OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: February 3, 2000, 14:17:29 ; Search time 49779.6 Seconds
(without alignments)
-75.331 Million cell updates/sec

Title: US-08-962-560A-3
Perfect score: 1235
Sequence: 1 CGAGGCTCAAGCTCCGGGGCG..........AGAAAAAAAAAAAAAAAAA 1235

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 821193 seqs, -1518192014 residues
Total number of hits satisfying chosen parameters: 1642386
Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :   GenEmbl:*

1: gb_bal:*	17: em_fun:*	33: gb_htg2:*
2: gb_ba2:*	18: em_hum1:*	34: gb_in1:*
3: gb_om:*	19: em_hum2:*	35: gb_in2:*
4: gb_ov:*	20: em_in:*	36: gb_bal:*
5: gb_pat:*	21: em_om:*	37: em_ba1:*
6: gb_ph:*	22: em_or:*	38: em_ba2:*
7: gb_pll:*	23: em_ov:*	39: em_hum3:*
8: gb_pl2:*	24: em_pat:*	40: em_hum4:*
9: gb_prl:*	25: em_ph:*	41: gb_pr4:*
10: gb_pr2:*	26: em_pl:*	42: gb_htg3:*
11: gb_pr3:*	27: em_ro:*	43: gb_htg4:*
12: gb_ro:*	28: em_sts:*	44: gb_htg5:*
13: gb_sts:*	29: em_sy:*	44: gb_htg6:*
14: gb_sy:*	30: em_un:*	45: gb_htg7:*
15: gb_un:*	31: em_vi:*	46: em_htg1:*
16: gb_vl:*	32: gb_htg1:*	47: em_htg2:*
		48: em_htg3:*
		49: em_hum5:*
		50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%Query Match	Length	DB	DB ID	Description
1	1218	98.6	13812	12	MMPRMGNS	Z47352 M.musculus
2	1119.2	90.6	13185	12	MMU88325	U88325 Mus musculu
3	1117.2	90.5	11193	12	MMU88325	U88325 Mus musculu
4	1098.4	88.9	11177	12	AF120490	AF120490 Mus muscu
5	1086	87.9	13187	12	AF180302	AF180302 Mus muscu
6	1051.2	85.1	1056	12	BTPRMTNP2	246939 R.norvegicu
7	904.2	73.2	909	12	AB000677	AB000677 Mus muscu
8	725.6	58.8	12589	40	AB000710	AB000710 Mus muscu
9 c	725.6	58.8	25417	32	AC002286	AC002286 Homo sapi
10	723.6	58.6	175950	33	HSAC002047	AC002047 Homo sapi
11	718.6	58.2	1216	9	AC007220	AC007220 Homo sapi
12	690.2	55.9	1094	10	AB000734	AB000734 Homo sapi
13	682.4	55.3	1082	9	HSU88326	U88326 Human suppr
14	623.2	50.5	908	40	AB000676	AB000676 Homo sapi
15	571.6	46.3	11270	41	AF132440	AF132440 Homo sapi
16	473.4	38.3	850	9	AC009140	AC009140 Homo sapi
17	101.4	8.2	678	9	AB004904	AB004904 Homo sapi
18	95.6	7.7	678	9	AB006967	AB006967 Homo sapi
19	94	7.6	682	40	AF159854	AF159854 Homo sapi
20	93.2	7.5	2792	12	RNO249240	AJ249240 Rattus no
21	91.4	7.4	2187	12	MMU88328	U88328 Mus musculu
22	90.8	7.4	863	12	AF075383	AF075383 Rattus no
23	82.2	6.7	1262	4	AF167297	AF167297 Gallus ga
24	69.6	5.6	774	5	AR035829	AR035829 Sequence
25	69.6	5.6	774	5	AR035856	AR035856 Sequence
26	69.6	5.6	1960	5	AR035809	AR035809 Sequence
27	69.6	5.6	1960	5	AR035836	AR035836 Sequence
28	69.6	5.6	2198	40	AF132297	AF132297 Homo sapi
29 c	64	5.2	181634	44	AC016253	AC016253 Homo sapi
30	62.2	5.1	22861	11	HSLUCA2	Z77852 Human DNA s
31	61	5.0	57960	44	AC016188	AC016188 Homo sapi
32 c	61	4.9	57960	32	CEY3986_3	CEY3986_3 Continuation (4 of
33	60.4	4.9	171663	43	HSJ1167H4	AL121914 Homo sapi
34	60.4	4.8	1737721	41	AC015866	AC015866 Homo sapi
35	59.6	4.8	108064	41	AC010404	AC010404 Homo sapi
36	59.4	4.8	110000	45	AC017014_1	AC017014 Continuation (2 of
37	58.8	4.8	148754	44	AC013648	AC013648 Homo sapi
38	58.4	4.7	46108	41	AC011100	AC011100 Homo sapi
39	58.2	4.7	4839	9	HUMOPIODRE	HUMOPIODRE Human putat
40	58.2	4.7	107475	45	AC017106	AC017106 Homo sapi
41	58	4.7	64096	44	AC015468	AC015468 Homo sapi
42 c	57.8	4.7	840	8	CNS01BNS	AL114464 Botrytis
43	57.6	4.7	181634	44	AC016253	AC016253 Homo sapi
44	57.4	4.6	1964	2	SC2C1A	AL133219 Streptomy
45	57.4	4.6	77671	43	AC013505	AC013505 Homo sapi

ALIGNMENTS

RESULT 1
MMPRMGNS
LOCUS   MMPRMGNS   13812 bp   DNA   ROD   11-DEC-1996
DEFINITION M.musculus PRM1, PRM2, PRM3 and TNP1 genes.
ACCESSION   Z47352
VERSION   Z47352.1 GI:1360004
KEYWORDS   PRM1 gene; PRM2 gene; PRM3 gene; protamine; TNP2 gene; transition

```
SOURCE          house mouse.
  ORGANISM      Mus musculus
                Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
                Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 13812)
  AUTHORS      Schlueter,G. and Engel,W.
  TITLE        Genomic sequence of the mouse protamine gene cluster
  JOURNAL      Unpublished
REFERENCE   2  (bases 1 to 13812)
  AUTHORS      Schlueter,G.
  TITLE        Direct Submission
  JOURNAL      Submitted (06-JAN-1995) Schlueter G., Institut fuer Humangenetik,
               Universitaet Goettingen, Gosslerstr.12d, Goettingen, germany, 37073
  REMARK       Revised by [3]
REFERENCE   3  (bases 1 to 13812)
  AUTHORS      Schlueter,G.
  TITLE        Direct Submission
  JOURNAL      Submitted (30-MAY-1996) Schlueter G., Institut fuer Humangenetik,
               Universitaet Goettingen, Gosslerstr.12d, Goettingen, germany, 37073
REFERENCE   4  (bases 1 to 13812)
  AUTHORS      Schlueter,G., Celik,A., Obata,R., Schlicker,M., Hofferbert,S.,
               Schlung,A., Adham,I.M. and Engel,W.
  TITLE
  JOURNAL      Mol. Reprod. Dev. 43 (1), 1-6 (1996)
  MEDLINE      96341725
COMMENT        On Jun 5, 1996 this sequence version replaced gi:886891.
               Related sequences: X07625, X07626, M6254.
               Sequence analysis of the conserved protamine gene cluster shows
               that it contains a fourth expressed gene
FEATURES             Location/Qualifiers
     source          1..13812
                     /organism="Mus musculus"
                     /strain="c129"
                     /db_xref="taxon:10090"
                     /chromosome="16"
     gene            653..902
                     /gene="Prm1"
     CDS             join(653..758,853..902)
                     /gene="Prm1"
                     /note="male germcell specific, haploid expressed;
                     chromatin binding"
                     /codon_start=1
                     /product="protamine 1"
                     /protein_id="CAA87410.1"
                     /db_xref="GI:1360005"
                     /db_xref="SWISS-PROT:P02319"
                     /translation="MARYRCCRSKSRSRCRRRRRRCRRRRRRCCRRRSY"
     gene            5389..5817
                     /gene="Prm2"
     CDS             join(5389..5647,5753..5817)
                     /gene="Prm2"
                     /note="Description: male germcell specific, haploid
                     expressed; chromatin binding"
                     /codon_start=1
                     /product="protamine 2"
                     /protein_id="CAA87411.1"
                     /db_xref="GI:1360006"
                     /db_xref="SWISS-PROT:P07978"
                     /translation="MVRYRMRSPSESPHQGPGQDHEREEQGQGQGLSPERVEDYGRTH
                     RGHHHRHRRCSRKRLHRIHRRRSCRRRRRHSCRRRRRHRGCRRSRRRRRCRCRKC
                     RHHH"
     gene            6498..6806
                     /gene="Prm3"
     CDS             6498..6806
                     /gene="Prm3"
                     /note="male germcell specific, haploid expressed"
                     /codon_start=1
                     /product="protamine 3"
                     /protein_id="CAA87412.1"
                     /db_xref="GI:1360007"
                     /db_xref="SWISS-PROT:Q62100"
                     /translation="MSSRCSKLSTGHEQPAQNTGHNRGHESSMKKLVACVSQDNFSLS
     gene            8786..9310
                     /gene="Tnp2"
     CDS             join(8786..9122,9294..9310)
                     /gene="Tnp2"
                     /note="male germcell specific, haploid expressed; chromatin
                     binding"
                     /codon_start=1
                     /product="transition protein 2"
                     /protein_id="CAA87413.1"
                     /db_xref="GI:1360008"
                     /db_xref="SWISS-PROT:P11378"
                     /translation="MDTKMQSLPTTHPHPHSSSSRPQSHTSNQCNQCTCSHHCRSCSQA
                     GHAGSSSPSPSGPPMKHPKPSVHSRHSPARPSHRGSCPKNRRTFEGKVSKRRAVRRR
                     RTHRAKRRTSGRRYK"
                     SEGEEEEEEEEEEEEEEEEEEEEEEEEQIPVKGKLLLEPEKQESAEDGEAQPSPEPRQQTHS
                     "
BASE COUNT     3400 a   3514 c   3869 g   3029 t
ORIGIN
```

Query Match
Best Local Similarity   98.6%;   Score 1218;   DB 12;   Length 13812;
Matches 1221; Conservative   0; Mismatches   5; Indels   0; Gaps   0;
Pred. No. 2.6e-210;

```
Qy      1 CGAGGCTCAAGCTCCGGGCGGATTCTGCGTGCCGCTCCTTGGGGTGTTGGCC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12356 CGAGGCTCAAGCTCCGGGCGGATTCTGCGTGCCGCTCCTTGGGGTGTTGGCC 12415

Qy     61 GGCCTGGTGCCACCCGGACGCGCCCGGGCTCACTGCCTCTGTTCTCCCCCATCAGCCGCCCG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12416 GGCCTGGTGCCACCCGGACGCGCCCGGGCTCACTGCCTCTGTTCTCCCCCATCAGCCGCCCG 12475

Qy    121 GACGGCTATGGCCCACCCCTCCAGCTGGCCCCTCGAGTAGGATGGTAGCCACGGCAACCAGGT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12476 GACGGCTATGGGCCCACCCCTCCAGCTGCGCCCCTCGAGTAGGATGGTAGCCACGGCAACCAGGT 12535

Qy    181 GGCAGCCGACACAATGCGATCTCCCGGCACGAGCGGCGGTCAGACCCCTCTC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12536 GGCAGCCGACACAATGCGATCTCCCGGCACGAGCGGCGGTCAGACCCCTCTC 12595

Qy    241 GTCCTCGTCTTCGTTCCTCGCCAGGCGCCCCCGTGCGTGCCGGGCCCCGGTCCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12596 GTCCTCGTCTTCGTTCCTCGCCAGGCGCCCCCGTGCGTGCCGGGCCCCGGTCCC 12655

Qy    301 AGCCCCAGCCCCTGGGACACTCACTTCCGCACCTTCCGCGATTACCGGCG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  12656 AGCCCCAGCCCCTGGGACACTCACTTCCGCACCTTCCGCGATTACCGGCG 12715

Qy    361 CATCACCGGGACCAGCGCGCGCTCCTGGACGCCGTGCGGCTTCTATTGGGG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  12716 CATCACCGGGACCAGCGCGCGCTCCTGGACGCCGTGCGGCTTCTATTGGGG 12775

Qy    421 GCACGGGGGCGGCACGAGCGGTGCGTGCCGCGGGGCGACCTTCTTGGTCGCGACAG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12776 GCACGGGGGCGGCACGAGCGGTGCGTGCCGCGGGGCGACCTTCTTGGTCGCGACAG 12835

Qy    481 TCGTCAACGGAACTGCTTCTTCGCGCTCAGCGTGAAGAATGGCTTCGGGGCCCACGAGCAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12836 TCGCCAACGGAACTGCTTCTTCGCGCTCAGCGTGAAGAATGGCTTCGGGGCCCACGAGCAT 12895

Qy    541 CCGCGTGCACTTCCAGGCCGGCGCCGCTTCCACTTGGACGACCTTCGACTG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  12896 CCGCGTGCACTTCCAGGCCGGCGCCGCTTCCACTTGGACGACCTTCGACTG 12955

Qy    601 CCTTTTCGGAGCTGCTGGAGCACTACGTGCGGCCGCATGTTGGGGCGCCCGCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  12956 CCTTTTCGGAGCTGCTGGAGCACTACGTGGCCGCATGTTGGGGCGCCCGCT 13015

Qy    661 GCGGCCAGCGCGCCGGCGCCGCTGCGCGCCAGCGAGCTGTGTTCGCGCCGGT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  13016 GCGGCCAGCGCGCCGGCGCCGCTGCGCGCCAGCGAGCTGTGTTCGCGCCGGT 13075

Qy    721 GGGTCGCGGAGAACCTGCGCGCGCGCGCATCCCTCTTTAAACCGGTACTCCGTGACTACCTGAGTTC 780
```

RESULT 2

```
LOCUS       MMU88325     1185 bp    mRNA            ROD       06-JUL-1997
DEFINITION  Mus musculus suppressor of cytokine signalling-1 (SOCS-1) mRNA,
            complete cds.
ACCESSION   U88325
VERSION     U88325.1  GI:2245381
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1185)
  AUTHORS   Starr,R., Willson,T.A., Viney,E.M., Murray,L.J.L., Rayner,J.R.,
            Jenkins,B.J., Gonda,T.J., Alexander,W.S., Metcalf,D., Nicola,N.A.
            and Hilton,D.J.
  TITLE     A family of cytokine-inducible inhibitors of signalling
  JOURNAL   Nature 387 (6636), 917-921 (1997)
  MEDLINE   97345633
REFERENCE   2  (bases 1 to 1185)
  AUTHORS   Starr,R., Willson,T.A., Viney,E.M., Murray,L.J.L., Rayner,J.R.,
            Jenkins,B.J., Gonda,T.J., Alexander,W.S., Metcalf,D., Nicola,N.A.
            and Hilton,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-1997) Cancer Research Unit, The Walter and Eliza
            Hall Institute of Medical Research, PO Royal Melbourne Hospital,
            VIC 3050, Australia
FEATURES             Location/Qualifiers
     source          1..1185
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="16"
                     /tissue_type="thymus"
     gene            119..757
                     /gene="SOCS-1"
     CDS             119..757
                     /gene="SOCS-1"
                     /codon_start=1
                     /product="suppressor of cytokine signalling-1"
                     /protein_id="AAB62400.1"
                     /db_xref="GI:2245382"
                     /translation="MVARNQVAADNAISPAAEPRRRSEPSSSSSSSPAAPVRPRPCP
                     AVPAPAPGDTHRHFRSHSDYRITRTSRALLDACGFYYGPLSVHGAHERLRAPVGTF
                     LVRDSRQRNCFFALSVKWASGPTSIRVRHFQAGRFHLDGSRETFDCLFELLEHVAAPR
                     RMLGAPLRQRVRRPLQELCRQRIVAAVGRENLARIPLNPVLRDYLSSSFPPQI"
BASE COUNT      183 a    413 c    334 g    255 t
ORIGIN
```

Query Match 90.6%; Score 1119.2; DB 12; Length 1185;
Best Local Similarity 97.2%;
Matches 1139; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
Pred. No. 1.8e-192;

```
Qy     56 TGGCCGGGCCTGTGCCACCCGGACGGCGCCGGCTCACTGCCTGTCTCCCCATCAGCGCAG 115
          || |  ||  || |  |||| |  |   |  | |  || | | | | |||||||||||||
Db     14 TGCGCGACAGCGCCGCCCGGAGCCCCGGAGCCCCCGGAGCCCCCGGAGCCCCCATCAGCGCAG 73

Qy    116 CCCCGGACCGCTATGGCCCACCCCTTCCAGCTGGCCCCTCGAGTAGGATGGTAGCACGCAAC 175
          || |||| ||| |||| |  |||||||||||||||||||||||| ||| ||| | |  ||||
Db     74 CCCCGGACCGCTATGGCCCACCCCTTCCAGCTGGCCCCTCGAGTAGGATGGTAGCACGCAAC 133

Qy    176 CAGGTGGCAGCGCAGCACAATGCGATCTCCCCGGCAGCAGAGCCCCCGACGGGGTCAGAGCC 235
          ||||||||||||||| ||||| ||||||||||||||||||||||||||||||||||||||||
Db    134 CAGGTGGCAGCGCAGCACAATGCGATCTCCCCGGCAGCAGAGCCCCCGACGGGGTCAGAGCC 193

Qy    236 TCCTCGTCCTCGTCTTCGTCCTCGCCAGCGGCCCCGGCGGCCGTCCCGTGCGTCCCGGGCG 295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194 TCCTCGTCCTCGTCTTCGTCCTCGCCAGCGGCCCCGGCGGCCGTCCCGTGCGTCCCGGGCG 253

Qy    296 GTCCCAGCCCCACGCCCCCTGGCGACACTCACTTCCGCCACCTTCCGCCTCCCACTCCGATTAC 355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 GTCCCAGCCCCACGCCCCCTGGCGACACTCACTTCCGCCACCTTCCGCCTCCCACTCCGATTAC 313

Qy    356 CGGCGGCATCACGCGGGACCAGACGGCGCCGCTCCTTCGGACGGCGCTGGCTTCTATGGGGACCCCTG 415
          ||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    314 CGGCGGCACCACGCGGGACCAGCAGGCGCCGCTCCTTCGGACGGCGCTTTCTATGGGGACCCCTG 373

Qy    416 AGCGGTGCACGGGGCGGGCGGCCCACGAGCGGCGCTGCGTGCCGCGGCCCGGGCCACCTTCTTGGTGCGC 475
          |||||||||||||| |||| |||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    374 AGCGGTGCACGGGCGGGCGGCCCACGAGCGGCGCTGCGTGCCGCGGCCCGGGCCACCTTCTTGGTGCGC 433

Qy    476 GACAGTCGTCAACGGAGAACTGCTTCTTTCCGCGTCAGCGTGCGAAGATGGGTTCGGGCCCCACG 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 GACAGTCGCCAACGGAGAACTGCTTCTTCGGCGTCAGCGTGAAGATGGGTTCGGGCCCCACG 493

Qy    536 AGCATCCCGGTGCACTTCCAGGCCGCCGCTTCCACTTGGACGGCGCCAGCCGCGAGACCTTC 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    494 AGCATCCCGGTGCACTTCCAGGCCGCCGCTTCCACTTGGACGGCGCCAGCCGCGAGACCTTC 553

Qy    596 GACTGCCTTTTCGAGCTGCTGGAGCACTACGTGGCGGCGCCGCGCCGCATGTTGGGGGCC 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    554 GACTGCCTTTTCGAGCTGCTGGAGCACTACGTGGCGGCGCCGCGCCGCATGTTGGGGGCC 613

Qy    656 CCCCTGCCCCAGCGGCGCGTGCGGCGGCCGCTGCAGGAGCTGTGTCGCCAGCGCATCGTGGCC 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    614 CCGCTGCGCCAGCGCGTGCGGCGGCCGCTGCAGGAGCTGTGTCGCCAGCGCATCGTGGCC 673

Qy    716 GCCGGTGGGGAGAACCTGGCGCGGATACCCCTCTTAACCCCGTGACTCCGGTGACTACCTG 775
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    674 GCCGGTGGGGAGAACCTGGCGCGGATACCCCTCTTAACCCCGTGACTCCGGTGACTACCTG 733

Qy    776 AGTTCCTTCCCCTTCCAGATCTGACCGGCGTGCCGCAGCATTAAGTGGGGGCG 835
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    734 AGTTCCTTCCCCTTCCAGATCTGACCGGCGTGCCGCAGCATTAAGTGGGGGCG 793

Qy    836 CCTTATTATTTCTTATTATTAATTAATTATTAATTTCTGGAACCACGTGGGAGCCCCTCC 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    794 CCTTATTATTTCTTATTATTAATTAATTATTAATTTTCTGGAACCACGTGGGAGCCCCTCC 853

Qy    896 CGGCCTGGGGTCGGAGGAGTGGTTGTGGAGGGTGAGATGCGCTCCCACTTCTGGCTGGAGAC 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    854 CGGCCTGGGGTCGGAGGAGTGGTTGTGGAGGGTGAGATGCGCTCCCACTTCTGGCTGGAGAC 913
```

Far-left alignment block:

```
Db  13076 GGGTCGGAGGAGTGGTTGTGTGAGGGTGAGATGCGCTCCCACTTCTGGCTGGAGACCTCAT 13135
Qy    781 CTTCCCCTTCCAGATCTGACCGGCGCTGCGCCGGCCCCAGCATTAAGTGGGGGCGCGCCTTA 840
Db  13136 CTTCCCCTTCCAGATCTGACCGGCGCTGCGCCGGCCCCAGCATTAAGTGGGGGCGCGCCTTA 13195
Qy    841 TTATTTCTTATTATTAATTAATTATTAATTTTCTGGAACCACGTGGGAGCCCCTCCCGGCCT 900
Db  13196 TTATTTCTTATTATTAATTAATTATTAATTTTCTGGAACCACGTGGGAGCCCCTCCCGGCCT 13255
Qy    901 GGGGTCGGAGGAGTGGTTGTGTGAGGGTGAGATGCGCCTCCCACTTCTGGCTGGAGACCTCAT 960
Db  13256 GGGGTCGGAGGAGTGGTTGTGTGAGGGTGAGATGCGCTCCCACTTCTGGCTGGAGACCTCAT 13315
Qy    961 CCCACCTCTCAGGGGTGGGGGTCTCCCCTCCCCTCCTCGGTCGTCCCCATCAGCGCAG 1020
Db  13316 CCCACCTCTCAGGGGTGGGGGTCTCCCCTCCCCTCCTCGGTCGTCCCCATCAGCGCAG 13375
Qy   1021 GTAGCAGCTTGTGTCTCTGGGGGCCAGGACCTGAATTCCACTCTCTCCCATGTTACAT 1080
Db  13376 GTAGCAGCTTGTGTCTCTGGGGGCCAGGACCTGAATTCCACTCTCTCCCATGTTACAT 13435
Qy   1081 ATTCCAGTATCTTTGCACAAACCAGGGGGTCGGGGAGGGTCTCTGGCTTCATTTTCTGC 1140
Db  13436 ATTCCAGTATCTTTGCACAAACCAGGGGGTCGGGGAGGGTCTCTGGCTTCATTTTCTGC 13495
Qy   1141 TGTGCAGAATACCTATTTTATATTTTTACACCCAGTTTAGGTAATAAAACTTTATTATGA 1200
Db  13496 TGTGCAGAATACCTATTTTATATTTTTACACCCAGTTTAGGTAATAAAACTTTATTATGA 13555
Qy   1201 AAGTTTTTTTTTTAAAAGAAAGAAAAAAA 1226
Db  13556 AAGTTTTTTTTTTAAAAAGAAAACAAAGA 13581
```

```
RESULT 3
AF120490    1193 bp    mRNA              ROD    10-AUG-1999
LOCUS       AF120490
DEFINITION  Mus musculus suppressor of cytokine signalling-1 (Socs1) mRNA,
            complete cds.
ACCESSION   AF120490
VERSION     AF120490.1  GI:4581880
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1193)
  AUTHORS   De Sepulveda,P., Okkenhaug,K., La Rose,J., Hawley,R.G., Dubreuil,P.
            and Rottapel,R.
  TITLE     Socs1 binds to multiple signalling proteins and suppresses steel
            factor-dependent proliferation
  JOURNAL   EMBO J. 18 (4), 904-915 (1999)
  MEDLINE   99146908
REFERENCE   2  (bases 1 to 1193)
  AUTHORS   De Sepulveda,P. and Rottapel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-1999) Experimental Therapeutics, Ontario Cancer
            Institute, 610 University Avenue, Toronto, ON M5G 2M9, Canada
FEATURES             Location/Qualifiers
     source          1..1193
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="16"
                     /cell_line="EML-C1"
     gene            1..1193
                     /gene="Socs1"
     CDS             130..768
                     /gene="Socs1"
                     /note="Socs-1"
                     /codon_start=1
                     /product="suppressor of cytokine signalling-1"
                     /protein_id="AAD24777.1"
                     /db_xref="GI:4581881"
                     /translation="MVARNQVAADNAISPAAEPRRSEPSSSSSSSPAAPVRVRRPCP
                     AVPAPAPGDTHFRFPRSHSDYRTRTRTSALLDAGFYWGPLSVGAHERLRAEPVGTF
                     LVRDSRQRNCFFALSVKMASGPTISRVHFQAGRFHLDGSRETFDCLFELLEHYVAAPR
                     RMLGAPLRQRRVRPLQELCRQRIVAAVGRENLARIPLNPVLRDYLSSFPFQI"
BASE COUNT      419 a    339 g    255 t    180 c
ORIGIN
```

Matches 1137;   Conservative   0;   Mismatches   32;   Indels   0;   Gaps   0;

```
Qy     56 TGGCCGGGACGGCTGTGCCACCCGGACGCCCGGCCTCACTGCCTCTGTCTCCCCATCAGCCAG 115
          ||||||| || | ||||||||||||| ||||||||||| ||||| |||||||||||||| ||
Db     25 TGGCCGGGACAGCGGCCCCCGGGAGCCCGGCCTCAGCGCGGGGTCCCCCCCGGCAGCGGCAG 84

Qy    116 CCCCGGACGGCTANTGGCCCACCCCTCCAGCTCGCCCCTCGAGTAGGATAGGTAGCACGCAAC 175
          ||||||||||| | ||||| |||||||||||||| ||||||| ||| ||| ||| |||||||
Db     85 CCCCGGACGGCTATGGGCCCACCCCTCCAGCTGGCCCCTCGAGTAGGATAGGTAGCACGCAAC 144

Qy    176 CAGGTGGCGGAGCCGGAGCAGCAATGCGGATCTCCCGGGCTCCCCGGCACCGGCGGTCGAGCGG 235
          ||||||||| |||||||||||| |||||| |||||| |||||||||||||||||||||||||
Db    145 CAGGTGGCGGAGCCGACAATGCGGATCTCCCGGGCAGCAGACCCCCGGCACCGGCGGTCGAGCCC 204

Qy    236 TCCTCGTCCTCGTCTTGTCTCGCCCAGCGCCCCGGCCGGGCCGGGCCCGGTCCCCGGCCCACG 295
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    205 TCCTCGTCCTCGTCTTCGTCTCGCCCAGCGCCCCGGCCGGGCCGGGCCCGGTCCCCGGCCCACG 264

Qy    296 GTCCCAGCGCCCAGCCCCTGCGACACTCACTTCCGCACCTTCCGCCTCCCACTCCGATTAC 355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 GTCCCAGCGCCCAGCCCCTGCGACACTCACTTCCGCACCTTCCGCCTCCCACTCCGATTAC 324

Qy    356 CGGGCCGCATCACCCGGACCAGGCCGCGCCTCCTGGACGCGCCTTCTATTGGGGACCCCTG 415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 CGGGCCGCATCACCCGGACCAGGCCGCGCCTCCTGGACGCGCCTTCTATTGGGGACCCCTG 384

Qy    416 AGGGTGCACGGGGCCGGCGCACGAGGGCCTGCGGTGCCGAGCGCCACCTTCTTGGGTGCGC 475
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    385 AGCGTGGCGAGCCGACAATGGCCGAGTCCCGGCAGCAGACCCCGGCACCGGCGGTCGAGCCC 444

Qy    476 GACAGTCGTCAACGGAGAACTGCTTCTTCGCGCTGCGCCGGGTGAAGAGTGGCGTTCGGGGCCCACG 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 GACAGTCGCCAACGGAGAACTGCTTCTTCGCGCTGTCGGTGAAGATGGCGTTCGGGGCCCACG 504

Qy    536 AGCATCCGGCGCACTTCCAGGCGCCGGCGCGGGTCCCACTTGGACGGCGCACCCGACACCTTC 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 AGCATCCGGCGCACTTCCAGGCGCCGGCGCGGTTCCACTTGGACGGCGCACCCGAGACCTTC 564

Qy    596 GACTGCCTTTTCGAGCTGCTGGAGCACTACGTGGCGGCGCCCCGCCCCCATGTTGGGGGCC 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    565 GACTGCCTTTTCGAGCTGCTGGAGCACTACGTGGCGGCGCCCCGCCCCATGTTGGGGGCC 624

Qy    656 CCGCTGCGCCAGCGGCGCGTGCGGCCCCTGCAGGAGCTGTGTCGCCAGCGCATCGTGGCC 715
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    625 CCGCTGCGCCAGCGGCGCGTGCGGCCCCTGCAGGAGCTGTGTCGCCAGCGCATCGTGGCC 684

Qy    716 GCGGTGGGTCGGGAGAACCTGGCGCGGATCCCTCTAAACCCGGTACTCCGGTGACTACCTG 775
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    685 GCCGTGGGTCGGGAGAACCTGGCGCGGCATCCCTCTTAAACCCGGTACTCCGGTGACTACCTG 744

Qy    776 AGTTCCTTCCCCTTCCAGATCTGACCGGCTGCCCGCGTGCCCCAGCAGCATTAAGTGGGGGCG 835
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    745 AGTTCCTTCCCCTTCCCAGATCTGACCGGCTGCCCGCGTGCCCCAGCAGCATTAAGTGGGGGCG 804

Qy    836 CCTTATTATTTCTTATTATTAATTATATTTTCTGGAACCACGTGGGAGCCCTCCC 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    805 CCTTATTATTTCTTATTATTAATTATATTTTCTGGAACCACGTGGGAGCCCTCC 864

Qy    896 CGCCTGGGGTCGAGGGAGTGGTTGTGTGGAGGGGTGAGAGAGCCCTCCCGGTCCCCGGGAC 955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    865 CGCCTGGGGTCGGAGGGAGTGGTTGTGTGGAGGGGTGAGAGAGCCCTCCCGGTCCCCGGGAC 924

Qy    956 CTCATCCCACCTCTCAGGGGTCGTCCTCCCTCCGGTGTCCTCCTCCCGGGTCCCC 1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    925 CTCATCCCACCTCTCAGGGGTGCTCCTCCCTCCCACTTCTGGCTGGAGATGCCTCCACCGGGTCCCC 984

Qy   1016 TGGTTGTAGCAGCTTGTGTCTGGGGCCCAGGAGGACCTGAATTCCAACTCCTACCCTCCATGTT 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    985 TGGTTGTAGCAGCTTGTGTCTGGGGGCCAGGAGGACCTGAATTCCAACTCCTACCCTCCATGTT 1044

Qy   1076 TACATATTCCCAGTACTTTTGCACAAACCAGGGGTCGGGGAGGGGTCTCTGGCTTCATTT 1135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 TACATATTCCCAGTACTTTTGCACAAACCAGGGGTCGGGGAGGGGTCTCTGGCTTCATTT 1104
```

(left column alignment)

```
Db    854 CGCCTCGGGTCCGGAGGGAGCTGGTGGTTGTGGGAGGGTGAGATGCCTCCCACTTCTGGCCTGGAGAC 913
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    956 CTCATCCCACCTCTCAGGGGTCGTCCTCCCTCCCTCCGGTGTCCTCCTCCCGGGTCCCCC 1015

Db    914 CTCATCCCACCTCTCAGGGGTGGGGGTGGGGTGCTCCCCCTCCCTCCTCGGGTCCTCCTCCCGGGTCCCCC 973
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1016 TGGTTGTAGCAGCTTGTGTCTGGGGGCCAGGAGGACCTGAATTCCAACTCCTACCCTCCATGTT 1075

Db    974 TGGTTGTAGCAGCTTGTGTCTGGGGGCCAGGAGGACCTGAATTCCAACTCCTACCCTCCATGTT 1033
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1076 TACATATTCCCAGTATCTTTGCACAAACCAGGGGTCGGGGAGGGGTCTCTGGCTTCATTT 1135

Db   1034 TACATATTCCCAGTATCTTTGCACAAACCAGGGGTCGGGGAGGGGTCTCTGGCTTCATTT 1093

Qy   1136 TCTGCGTGTGCAGAATATCCTATTTATTTATATTTTACAGCCAGTTTAATTAATAAACTTTAT 1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1094 TCTGCGTGTGCAGAATATCCTATTTATTTATATTTTACAGCCAGTTTAGTAATAAACTTTAT 1153

Qy   1196 TATGAAAGTTTTTTTTTTTAAAAGAAAAAAAAAA 1227
          ||||||||||||||||||||||||||||||||||
Db   1154 TATGAAAGTTTTTTTTTTTTAAAAAAAAAAAAA 1185
```

```
Qy   1136 TCTTGCTGTGCAGAGAATATCCTATTATTTATATTTTTACAGCCAGTTTAGGTAATAAAACTTTAT 1195
          |||||||||| ||||||||||||| ||||||||| |||||||||||||||||||||||||||||||
Db   1105 TCTTGCTGTGCAGAATATCCTATTTATTATTTATTTTTACAGCCAGTTTAGGTAATAAAACTTTAT 1164

Qy   1196 TATGAAAGTTTTTTTTTTAAAAGAAAAAA 1224
          ||||||||||||||||||||||||| |||
Db   1165 TATGAAAGTTTTTTTTTTAAAAGAAACAAA 1193
```

RESULT 4

```
LOCUS       AF180302    1177 bp    mRNA           ROD       06-SEP-1999
DEFINITION  Mus musculus suppressor of cytokine signalling-1 (Socs1) mRNA,
            complete cds.
ACCESSION   AF180302
VERSION     AF180302.1  GI:5825619
KEYWORDS    .
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1177)
AUTHORS     Gregorieff,A., Pyronnet,S., Sonenberg,N. and Veillette,A.
TITLE       Regulation of suppressor of cytokine signaling-1 (SOCS-1)
            expression by translational repression
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 1177)
AUTHORS     Gregorieff,A., Pyronnet,S., Sonenberg,N. and Veillette,A.
TITLE       Direct Submission
JOURNAL     Submitted (24-AUG-1999) McGill Cancer Centre, McGill University,
            3655 Drummond Street, Montreal, Quebec H3G 1Y6, Canada
FEATURES             Location/Qualifiers
     source          1..1177
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="16"
     gene            1..1177
                     /gene="Socs1"
     CDS             135..773
                     /gene="Socs1"
                     /note="SOCS-1"
                     /codon_start=1
                     /product="suppressor of cytokine signalling-1"
                     /protein_id="AAD53324.1"
                     /db_xref="GI:5825620"
                     /translation="MVARNQVAADNAIASPAAEPRRSEPSSSSSSSPAAPVRPRPCP
                     AVPAPAGDTHFRIFRSHSDYRIRTRTSALLDACFYWGPLSVHGAHERLRAEPVGTF
                     LVRDSRORNCFFALSVKMASGPTSIRVHFQAGRFHLDGSREFFDCLFELLEHYVAAPR
                     RMLGAPLRQRRVRPLQELCRQRIVAAVGRENTLARIPLNPVLRDYLSSFFPYQI"
BASE COUNT      171 a    420 c    340 g    246 t
ORIGIN
```

Best Local Match

```
Best Local Similarity    88.9%;    Score 1098.4;    DB 12;    Length 1177;
Matches 1117;   Conservative   0;   Pred. No. 9.8e-189;
Query Match 97.3%;   Mismatches   31;   Indels   0;   Gaps   0;
```

```
Qy     56 TGGCCGGCCTGTGCCACCCGAGACCCCGGCCTCACTGCCTCGTTCTCCCCATACGCCAG 115
          ||| || |||||||||||||||||||||| ||||  ||  || | ||||| | |||||
Db     30 TGGCCGACAGCCCCCCGGAGCCCCCGGCTCCCCGCGCTCTCCCGGCCAG 89

Qy    116 CCCCGGACGCGTATGGCCCCCTCCACCCTCCACTGCCTCGAGTAGGTGGCGGCGAAC 175
          |||||||||| || |||||| |||||| || ||||| |||||||||| |||| | |||
Db     90 CCCCGGACGCCTANGGCCCACCCCTCCAGCTGGCCCTCCGAGTAGGATGGGTAGCACGCAAC 149

Qy    176 CAGGTGGCGAGCCGACACAATGCGACTCTCCCGGCAGCAGGAGCCCCGGCAGTCGAGACGC 235
          ||||||| ||||||| ||  | |||||||||||||| || | ||||||| ||||||||||
Db    150 CAGGTGGGCAGCCGACGATCTCCCCGGCACGACAGCCCCGGGTCAGAGCCC 209

Qy    236 TCCTCGCGTCCTCGTCTTCGTTCGTCCTCGCGAAGGGCCCCGGGGCCCCGGGGCCG 295
          |||||| |||||||||||||| | |  |||  ||||| |||||||||| ||||| ||
Db    210 TCCTCGTCCTCGTCTTCGTCTCTGCCACGGCGGCCGGGGCCCTGCCCGGCGG 269

Qy    296 GTCCCAGCCCATCACGCGCCAGCCCAGCGCCTGGGGCGGCGGGACACTCACTTCCGCCGATTAC 355
          |||||||||||  ||||||||||||| |||||||| |||||| ||||||||  ||||||||||
Db    270 GTCCCAGCCCCCAGCGGCCAGCCCCTGGGCCGGCACTCACTTCCGCCGATTAC 329

Qy    356 CGGCGCCATCACGCGCGGACCCAGCCCCTGCCGATCACGCGCGCGCCTTCCTGAGCGCCCCTG 415
          ||||| ||||||||||| |||| ||||| || || |||||| || ||||| | || |||||||
Db    330 CGGCGCCATCACGCGGACCAGCGACGCTGCCCGCGCTCCTTCTTGGGGACCCCTG 389

Qy    416 AGCCGGTGCACGGGCCGCCACGAGCAGCGGGTGCGTGCCGAGCCCTGGGGCACCTTCTTGGTTGCGC 475
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||| || ||
Db    390 AGGCGTGCACGGGCCGCCACGAGCGGCGGGTGCGTGCCGAGCCCTGGGGCACCTTCTTGGTTGCGC 449

Qy    476 GACAGTCGTCAAACGGAGAACTGGTTCTTCTGCGCGTTCAGGGTGAAGATGGCTTCGGGCGCCACG 535
          || |||| |||||||||||||| |||||||||||| |||||||||| |||||| ||| ||||||||
Db    450 GACAGTCGCCAAACGGAGAACTGCTTCTTCTGCGCCTTCAGGGTGAAGATGGCTTCGGGCGCCACG 509

Qy    536 AGCATCCCGCGTGCCACTTCCAGCGCCCGGCGGCCTTCCACTTGGACGGCGGCGGGGAGACCTTC 595
          |||||||||||||||||||||||||||| || || |||||||||||||||||  |||||||||||
Db    510 AGCATCCCGCGGGCCACTTCCAGGCGCCGGTGGCCTTCCACTTGGACGGCGGCGGGGAGACCTTC 569

Qy    596 GACTCGCCTTTCGACGTCGCTGGAGCACTACGTGCGGGGCGGGGCGCCGCGGGGCGGGGCC 655
          |||| |||||||||||| ||| ||||||||||||||||| || ||  | || |||||||||
Db    570 GACTGCCTTTCGACGTCGGTGGAGCACTACGTGGGGGGCCGGGGCCGCCGCGGGGCGGGGCC 629

Qy    656 CCGCTGCGCCAGCGCCCGCGGCGTCGCGGGGCGCTGCAGGAGCGCGCGATCGGCGGC 715
          |||||||||||||||||  || || |||||||||||||||||||| ||| |||||||
Db    630 CCGCTGCGCCAGCGCCGCGGCGTCGTCGGCGCTGCAGGAGCCGCCGATCGGTGGGC 689

Qy    716 GCCGTGGGGTCGCGGGAAACTGGGCGGCCCATCCCTTCTTAAACCCGGTACTCCGTGACTACCTG 775
          ||||||||||||||||| || ||||| ||||||||||||||||||||||||||||||||||||||
Db    690 GCCGTGGGGTCGCGGGAGAACCTGGGCCGCCCATCCCTTCTTAAACCCGGTACTCCGTGACTACCTG 749

Qy    776 AGTTCCTTCCCCTTCCCAGATCTTGACGGTCGCGCTGCCGCAGCGCAGCAATTAAGTGGGGCG 835
          ||||||||||||||||||||||| |||||| |||||||||||||||||||| ||||||||||
Db    750 AGTTCCTTCCCCTTCCCAGATCTCGACGGCGTCGCGCTGCCGCAGCGCAGCATTAAGTGGGGCG 809

Qy    836 CCTTATTATTTTCTTATTATTAATTATTTTAATAATTATTATTTTCTGGAACCACGTGGGGAGCCCTCC 895
          |||||||||| |||| ||||| ||||||||||||||||||| |||| |||| || |||| |||||||||
Db    810 CCTTATTATTTCTTATTATTAATTATTTTAATAATTATTATTTTTCTGGAACCACGTGGGGAGCCCTCC 869

Qy    896 CGCCTGGGGTCGGAGGGAGTGGGTTGTGGGAGGGAGAGTGAGACATGGCTGAGAC 955
          |||||||||||||||||||||||| ||||||||||||| |||| ||||| |||||
Db    870 CGCCTGGGGTCGGAGGGAGTGGGTTGTGTGGGAGGGAGAATGCTGAGAC 929

Qy    956 CTCATCCCACCTCTCAGGGGGTGGGGGTGCTCCCCCGGGTCGGGGTCCCC 1015
          |||| || |||||||||||||||||| ||||||||||||  ||||| ||||
Db    930 CTCACCCACCTCTCAGGGGTGCTCCCCGGGTCTCCCTGGTCTCCCTGGGGTCCCCCC 989

Qy   1016 TGGTTGTGTTAGCAGCTTGTGTCTGGGGGGCCAGGAAGTTCCACTCCTACCTCTCCAATGTT 1075
          |||| || |||||||||||||||||||||||||||| ||||||||| || || | ||||||
Db    990 TGGTTGTTAGCAGCTTGTGTGTCGGGGGCCAGGACCTGAATTCCACTCCTACCTCTCCAATGTT 1049

Qy   1076 TACAATATTCCCAGTATCTTTGCACAAAACAGGGGTGTCGGGGAGGGTCGCCTTCATTTT 1135
          ||||||||||||||||||| |||||||||| ||||| |  ||||||||||||||||||||
Db   1050 TACATATTCCCAGTATCTTTTGCACAAAACAGGGGTCGGGGAGGGTTCTCGGCCTTCATTTTT 1109

Qy   1136 TCTTGCTGTGCAGAGAATATCCTATTTATTTATTTTTACAGCCAGTTTAGGTAATAAAACTTTAT 1195
          |||||||||||||| ||||| |||||| |||||||||||||||| |||||||||||| |||||||
Db   1110 TCTCTGCTGTGCAGAATATCCTATTTTATTTATATTTTTACAGCCAGTTTAGGTAATAAAACTTTAT 1169

Qy   1196 TATGAAAG 1203
          ||||||||
Db   1170 TATGAAAG 1177
```

RESULT 5

```
LOCUS       BTPRMTNP2   13187 bp    DNA            ROD       11-DEC-1996
DEFINITION  R.norvegicus PRM1, PRM2, PRM3 and TNP2 genes.
ACCESSION   Z46939
VERSION     Z46939.1  GI:1355527
KEYWORDS    PRM1 gene; PRM2 gene; PRM3 gene; protamine; TNP2 gene; transition
```

```
SOURCE      protein 2.
            Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1 (bases 1 to 13187)
  AUTHORS   Schlueter,G., Celik,A., Obata,R., Schlicker,M., Hofferbert,S.,
            Schlung,A., Adham,I.M. and Engel,W.
  TITLE     Sequence analysis of the conserved protamine gene cluster shows
            that it contains a fourth expressed gene
  JOURNAL   Mol. Reprod. Dev. 43 (1), 1-6 (1996)
  MEDLINE   96341725
REFERENCE   2 (bases 1 to 13187)
  AUTHORS   Schlueter,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1994) Schlueter G., Institut fuer Humangenetik,
            Universitaet Goettingen, Gosslerstr.12d, Goettingen, FGR, 37073
REFERENCE   3 (bases 1 to 13187)
  AUTHORS   Schlueter,G.
  REMARK    Revised by [3]
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAY-1996) Schlueter G., Institut fuer Humangenetik,
            Universitaet Goettingen, Gosslerstr.12d, Goettingen, FGR, 37073
REFERENCE   4 (bases 1 to 13187)
  AUTHORS   Schlueter,G., Celik,A., Obata,R., Schlicker,M., Hofferbert,S.,
            Schlung,A., Adham,I.M. and Engel,W.
  TITLE     Sequence analysis of the conserved protamine gene cluster shows
            that it contains a fourth expressed gene
  JOURNAL   Unpublished
COMMENT     On Jun 5, 1996 this sequence version replaced gi:886853.
FEATURES             Location/Qualifiers
     source          1..13187
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /chromosome="10"
     gene            829..1075
                     /gene="PRM1"
     CDS             join(829..934,1026..1075)
                     /gene="PRM1"
                     /note="haploid expressed, male germcell specific,
                     chromatin binding"
                     /codon_start=1
                     /product="protamine 1"
                     /protein_id="CAA87061.1"
                     /db_xref="GI:1355529"
                     /db_xref="SWISS-PROT:P10118"
                     /translation="MARYRCRSKSRSRCRRRRRRCRRRRRRRCRRRRSY
                     TFRCKRY"
     repeat_unit     1800..1860
                     /note="ID element"
     repeat_unit     complement(2270..2330)
                     /note="ID element"
     repeat_unit     complement(2405..2580)
                     /note="B2 element"
     gene            5326..5760
                     /gene="Prm2"
     CDS             join(5326..5586,5695..5760)
                     /gene="Prm2"
                     /note="haploid expressed, male germcell specific,
                     chromatin binding"
                     /codon_start=1
                     /product="protamine 2"
                     /protein_id="CAA87062.1"
                     /db_xref="GI:1355529"
                     /db_xref="SWISS-PROT:P11248"
                     /translation="MVRYRMRPSEGQHQFPGQDHREEQGQGQELSPERVEDYGRTE
                     RGHHHRRCKRLHRIHKRRSCRRRRHSCHRRRRHRRRVSTQGCRRSRRRSCRRK
                     CRWHYY"
     gene            6513..6827
                     /gene="Prm3"
     CDS             6513..6827
                     /gene="Prm3"
                     /note="haploid expressed, male germcell specific"
                     /codon_start=1
                     /product="protamine 3"
                     /protein_id="CAA87063.1"
                     /db_xref="GI:1359930"
                     /translation="MGSRCAKLSTGHGPAQNTGHSRGHEBSSMKKLVACVSQDNFSLSS
                     EGEEEEDEEDEEDEEDDDEEDDEEDEEEQIPVKGLKLLLEPEKQDGAEDAVAQPSEPKQR
                     HS"
     gene            8710..9244
                     /gene="Tnp2"
     CDS             join(8710..9037,9225..9244)
                     /gene="Tnp2"
                     /note="haploid expressed, male germcell specific,
                     chromatin binding"
                     /codon_start=1
                     /product="transition protein 2"
                     /protein_id="CAA87064.1"
                     /db_xref="GI:1359931"
                     /db_xref="SPRREMBL:Q64390"
                     /translation="MDIKMQSLPTTHHPHPHSSSRPQSHTNNQCACSHHCRSCSQAGHP
                     SSSSPSPGPPTRHPKTPMHSRYSPSRPSHRGSCPKNRKTLEGRVSRKKAVRRKRTH
                     RAKRRSSVGRRYX"
BASE COUNT     3174 a   3484 c   3770 g   2759 t
ORIGIN
```

```
Query Match         87.9%;  Score 1086;  DB 12;  Length 13187;
Best Local Similarity  94.5%;  Pred. No. 1.6e-186;
Matches 1160;  Conservative  0;  Mismatches  60;  Indels  8;  Gaps  3;
```

```
Qy       1 CGAGGCTCAAGCTCCGGGCGGGATTCTGCGTGCCGCTCTCGCTCCTTGGGGTCGTGTTGCC  60
           |||||||||| |||||| ||||||| ||||||||||||| || ||||||||||||||||||
Db   11959 CGAGGCTCAGGCTCCGGGCGGGGATTCTGCGTGTCACCCTCGCTCCTTGGGGTCGCTTGCC  12018

Qy      61 GGCCTGTGCCACCCGGACGGCCCCGGCTCACTGCCTGCTCTGTCTCCCCATCAGCCAGCCCG  120
           |||||||||||||||||||||||| |||||| |||||||| |||||| |||||||||||||
Db   12019 GGCCTGTGCCACCCGGACGGCCCCGGCTCACTGCCTGCTCTGTCTCCCCATCAGCCAGCCCG  12078

Qy     121 GACGCTAYGGCCCACCCCTTCCAGCGTGGGCCCCTCGAGTAGGAGTTAGCAACCAGGT  180
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   12079 GACGCTAYGGCCCACCCCTTCCAGCGTGGGCCCCTCGAGTAGGAGTTAGCAACCAGGT  12138

Qy     181 GGCAGCCGACAATGCGATTCTCCCGGCACCAGAGACCCCGACGGCGGGTCAGAGCCCTCTC  240
           |||||||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||
Db   12139 GGAAGCCGACAAATGCGAATCTCCCGGCATCAGAGCCCCGACGGCGGGTCAGAGCCCAATCCTC  12198

Qy     241 GTCCTCGTCTTCTGTCCCTGCCACGGCCCCCGGTGCGTCCCCCGCGTCCCCGCGGGTCCC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12199 GTCCTCGTCTTCTGTCCCTGCCACGGCCCCCGGTGCGTCCCCCGCGTCCCCGCGGGTCCC  12258

Qy     301 AGCCCCAGCCCCTGGGCGACACTCACTTCCGCACCTTCCGGATTACCGGCG  360
           ||||||||||||||| |||||||||||||| ||||||||||||||||||||
Db   12259 GGGCCCCGGCTCCGGGCGACACTCACTTCCGCACCTTCCGGATTACCGGCG  12318

Qy     361 CATCACGCGGACCAGGCGGCCGCTCCTGGAGGCCCTGGCGCTTCTACTGGGGGACCCCTGAGGCT  420
           ||||| ||||||||||||| ||||||||||||||||||||| ||||||||||||||||||||||
Db   12319 CATCACGCGGACCAGGCGGCCGCTCCTGGAGGCCCTGGCGCTTCTACTGGGGGACCCCTGAGGGT  12378

Qy     421 GCACGGGCGCGCACGAGCGGCTGCGTGCGCGAGCCCGTGGGCCACCTTGCTGCGGCGACAG  480
           || ||||||| |||| || ||||||||||||||||||||||||| || |||||||||||||
Db   12379 GCAGGGGCGGCGCACGAACGGCTGCGTTCCGAACCCGTGGGCCACCTTCTTGGTGCGGCGACAG  12438

Qy     481 TCGTCAACGGAACTGCTTCTTCGGCGTCAGCGTGAAGAGATGGCTTCGGCCCCACAGCAT  540
           ||| ||||||||| |||||||||||||||| || |||||||||||||||||||||||||
Db   12439 TCGCCAGCGGGAACTGCTTCTTCGCGCTCAGCGTGAAGATGGCTTCGGCCCCACAGCAT  12498

Qy     541 CCGCGTGCACTTCCAGGCGCCGCCCTTCCACTTGGACGGCAGCCCGCGAGACCTTCGACTG  600
           ||| ||||||||||||| ||||||||||||||| || ||||||||| |||||||||||||
Db   12499 TCGTGTGCACTTCCAGGCGCCGCCCTTCCACCTGGACGGCAACCGCGAGACCTTCGACTG  12558

Qy     601 CCTTTTCGAGCTCGCTGGAGACCACTACGTGGGCGGCGGCGCCGCCGCCGCCCCGCGGCGT  660
           |||||||||||||| | ||||||||||||||||||||||||||||||||||||||||||
Db   12559 CCTCTTCGAGCTCGCTGGAGCACTACGTGGGCGGCGGCGCCGCCGCCGCCCCGCGCACT  12618
```

RESULT 6

```
LOCUS       AB000677     1056 bp    mRNA            ROD       22-JUL-1997
DEFINITION  Mus musculus mRNA for JAB, complete cds.
ACCESSION   AB000677
VERSION     AB000677.1  GI:2274771
KEYWORDS    JAB.
SOURCE      Mus musculus
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1056)
  AUTHORS   Yoshimura,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JAN-1997) to the DDBJ/EMBL/GenBank databases. Akihiko
            Yoshimura, Kurume university, Institute of Life Science; 2432-3
            Aikawa-machi, Kurume 839, Japan (Tel:0942-37-6313,
            Fax:0942-31-5212)
REFERENCE   2  (sites)
  AUTHORS   Yoshimura,A.
  TITLE     Cloning of a novel JAK kinase inhibitory SH2 protein
  JOURNAL   Unpublished (1997)
REFERENCE   3  (sites)
  AUTHORS   Endo,T.A., Masuhara,M., Yokouchi,M., Suzuki,R., Sakamoto,H.,
            Mitsui,K., Matsumoto,A., Tanimura,S., Ohtsubo,M., Misawa,H.,
            Miyazaki,T., Leonor,N., Taniguchi,T., Fujita,T., Kanakura,Y.,
            Komiya,S. and Yoshimura,A.
  TITLE     A new protein containing an SH2 domain that inhibits JAK kinases
  JOURNAL   Nature 387 (6636), 921-924 (1997)
  MEDLINE   97345634
FEATURES             Location/Qualifiers
     source          1..1056
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1..639
                     /codon_start=1
                     /product="JAB"
                     /protein_id="BAA21538.1"
                     /db_xref="GI:2274772"
                     /translation="MVARNQVAADNAISPAAEPRRSREPSSSSSSSSSSPPAPVRRRCP
                     GVPAPAPGDHFRTPRSHSDYRRIRTRSALLDACGFYWGPLSVHQAHERLRAEPVGTF
                     LVRDSRQRNCFFALSVKMASGPTSIRVHFQNGRFHLDGNRETFDCLFELLEHYVAAPR
                     RMLGAPLRQRRVRPLQELCRQRIVAAVGRENLARIPLNPVLRDYLSSFPQI"
BASE COUNT      159 a    354 c    299 g    244 t
ORIGIN
```

Query Match 85.1%; Score 1051.2; DB 12; Length 1056;
Best Local Similarity 99.7%; Pred. No. 3.1e-180;
Matches 1053; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy   161 ATGGTAGCACGGCAACGAAGGTGGGCACGCGACAAATGGGATCTCCCGCGCGCAGAGAGAGCCCGA 220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGGTAGCACGGCAACGAAGGTGGGCACGCGACAAATGGGATCTCCCGCGCGCAGAGAGAGCCCGA 60

Qy   221 CGGGCGGTCAGAGCCCTTCGGTCTTCGTCTTCGTCTTCGCTCCAGCCGCGGGGCCGGTGCGTCCC 280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CGGGCGGTCAGAGCCCTTCGGTCTTCGTCTTCGTCTTCGCTCCAGCCGCGGGGCCGGTGCGTCCC 120

Qy   281 CGGGCCCGTCGCCGCGGCCCGTCCAGCCCGACACTCACTCACTTCGCCGACCTTCCG 340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CGGGCCCGTCGCCGCGGCCCGTCCAGCCCGACACTCACTCACTTCGCCGACCTTCCG 180

Qy   341 TCCCACTCCGATTACCGGCGCATCACGCGGACCAGGAGCACTAGTGGGCGGGCTTC 400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TCCCACTCCGATTACCGGCGCATCACGCGGACCAGGAGCACTAGTGGGCGGGCTTC 240

Qy   401 TATTGGGGACCCCTGAGCGTGCACGGGCCGCACGAGCCGGGCCCGTGGGC 460
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TATTGGGGACCCCTGAGCGTGCACGGGCCGCACGAGCCGGGCCCGTGGGC 300

Qy   461 ACCTTCTTGGTGCGCGGACAGTCGCAACGGAACGCTTCTTCGCGGCTCAGCGTGAAGATG 520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACCTTCTTGGTGCGCGGACAGTCGCAACGGAACGCTTCTTCGCGGCTCAGCGTGAAGATG 360

Qy   521 GCTTCGGGCGCCCACGAGCAATCCGCGTGCACTTCCAGGCGCCGCGTTCCACTTGACGGC 580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCTTCGGGCGCCCACGAGCAATCCGCGTGCACTTCCAGGCGCCGCGTTCCACTTGACGGC 420

Qy   581 AGCCCGCGAGACCTTCGACTCGCCTTTTCGAGCTGCTGGAGCACTAGTGGGCGGGCCGC 640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AACCCGCGAGACCTTCGACTCGCCTTTTCGAGCTGCTGGAGCACTAGTGGGCGGGCCGC 480

Qy   641 CGCATGTTGGGGGGCCCCGCTGCGCCAGCCGGGCGTTGCCGGGCCGCAGGGAGCTGTGTCGC 700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CGCATGTTGGGGGGCCCCGCTGCGCCAGCCGGGCGTTGCCGGGCCGCAGGGAGCTGTGTCGC 540

Qy   701 CAGCGCCATCGTGGCCGCCCGTGGGTGCGAGAACCTGGCGGCGCCATCCCTCTTAACCCGGTA 760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CAGCGCCATCGTGGCCGCCCGTGGGTGCGAGAACCTGGCGGCGCCATCCCTCTTAACCCGGTA 600

Qy   761 CTCCGTGACTACCTGAGTTCCTTCCCAGATCTGACCGGCTGCCGCTGGCCGCTGCCGCAG 820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTCCGTGACTACCTGAGTTCCTTCCCAGATCTGACCGGCTGCCGCTGGCCGCTGCCGCAG 660

Qy   821 CATTAAGTAGTGGGGGGCGGCCTTAATTATTTCTTAATTATTTATTATTATTTTTCTGGAACCA 880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CATTAAGTAGTGGGGGGCGGCCTTAATTATTTCTTAATTATTTATTATTATTTTTCTGGAACCA 720

Qy   881 CGTGGGAGCGCCCTCCCGCGCCTGGGGTCGGAGGGGTAGGTGAGATGCCTCCCA 940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CGTGGGAGCGCCCTCCCGCGCCTGGGGTCGGAGGGGTAGGTGAGATGCCTCCCA 780

Qy   941 CTTCTGGCTGGAGACTCATCCCACCTCTCAGGGGGTGGGGTGCTCCCCCTCCCTGTGTGCTC 1000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CTTCTGGCTGGAGACTCATCCCACCTCTCAGGGGGTGGGGTGCTCCCCCTCCCTGTGTGCTC 840
```

```
Qy   661 GGGGTCGCGAGAACCTGGGCCGCCATCCCCGGTGCGCCGGCCAGGAGGCTGCTGCGGTC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 12619 GCGCCAGCGCGCCGGCTGCGCCGGGCTGCGCAGGAGGCTGTGTCGCGCAGCGCATCGTT 12678

Qy   721 GGGTCGGCGAGAACCTGGCGGCGCCATCCCCTCTTAACCCGGTACTCCGTGACTACTCGAGTTC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 12679 GGGTCGGCGAGAACCTGGCGGCGCCATCCCCTCTTAACCCGGTACTCCGTGACTACTCGAGTTC 12738

Qy   781 CTTCCCCGTTCCAGATCTGACCGGCTGTG-CCGCAGCATTAAGTAGTGGGGGGCGGCCTT 839
         |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db 12739 CTTCCCCAGATCTGACCGGCTGCCGCCGCAGCATTAAGTAGTGGGGGGCGGCCTT 12798

Qy   840 ATTATTTCTTAATTATTTATTATTATTTTTCTGGAACCACACGTGGGGGAGCCCTCCCGCC 899
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 12799 ATTATTTCTTAATTATTTATTATTATTTTTCTGGAACCACACGTGGGGGAGCCCTCCCGCC 12858

Qy   900 TGGGTCGGCGGAGAGTGGGTTGTGGAGGGTGAGAGTGGCCCTCCCAGCCCTCCAGCCCTCA 959
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 12859 TAGGTCGGACGGAGGAGTGGGTGTGTGGAGGGTGAGAGTGGCCCTCCCAGCCCTCCAGCCCTTA 12918

Qy   960 TCCCACCTCTCAGGGGGTGGGGTCGCTCCCCCCTCCGTCCCCTCCGGGTTCCCCCCTGGT 1019
         ||||||||| |||||||||||||||       |||||||||||||||| ||||||||||
Db 12919 TCCCACCTCTC-----GGGGGCCTCCCCCCTCTCGGTGCTGCTCCCCGGT-CCCCTGGT 12971

Qy  1020 TGTAGCAGCTTGTGTGTCTGGGGGGCCAGGAGACCTGAATTCCACTCCTACCTCCA 1079
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 12972 TGTAGCAGCTTGTGTGTCTGGGGGGCCAGGAGACCTGAATTCCACTCCTACCTCCA 13031

Qy  1080 TATTCCCAGTATCTTTGCACAAACCAGGGGTCTCCTGGGGGAGGGTCTCGGGGAGG 1139
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 13032 TGTTCCCAGTATCTTTGCACAAACCAGGGGTCTCCTGGGGGAGGGTCTCGGGGAGG 13091

Qy  1140 CTGTGCAGAATATCCTATTTTATATTTTTACAGCGCAGTTTAGGTAATAAAACTTTATTATG 1199
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 13092 CTGTGCAGAATATTCTATTTTACATCCAGTTTAGATAATAAAACTTTATTATG 13151

Qy  1200 AAAGTTTTTTTTTTTAAAAGAAAAAAAAA 1227
         |||||||||||||||||||||||||||||
Db 13152 AAAGTTTTTTTTTTTTTAAAGAAAACAAAGA 13179
```

```
RESULT 7

LOCUS       AB000710      909 bp    mRNA            ROD       22-JUL-1997
DEFINITION  Mus musculus mRNA for stat-induced stat inhibitor-1, complete cds.
ACCESSION   AB000710
VERSION     AB000710.1  GI:2274773
KEYWORDS    stat-induced stat inhibitor-1; SSI-1.
SOURCE      Mus musculus cDNA to mRNA, clone_lib:thymus cDNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   Kishimoto,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-1997) to the DDBJ/EMBL/GenBank databases.
            Tadamitsu Kishimoto, Osaka university, Medical school; Yamadaoka
            2-2, Suita 565, Japan (E-mail:matsumoto@imed3.med.osaka-u.ac.jp,
            Tel:+81-6-879-3831, Fax:+81-6-879-3839)
REFERENCE   2  (bases 1 to 909)
  AUTHORS   Naka,T., Narazaki,M., Hirata,M., Tomoshige,M., Minamoto,S.,
            Nishimoto,N., Kajita,T., Yoshizaki,K., Akira,S. and Kishimoto,T.
  TITLE     Structure and function of a novel STATs-induced inhibitor of STATs
            function-1(SIS-1)
  JOURNAL   Unpublished (1997)
REFERENCE   3  (sites)
  AUTHORS   Naka,T., Narazaki,M., Hirata,M., Matsumoto,T., Minamoto,S.,
            Aono,A., Nishimoto,N., Kajita,T., Taga,T., Yoshizaki,K., Akira,S.
            and Kishimoto,T.
  TITLE     Structure and function of a new STAT-induced STAT inhibitor
  JOURNAL   Nature 387 (6636), 924-929 (1997)
  MEDLINE   97345635
FEATURES             Location/Qualifiers
     source          1..909
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone_lib="thymus cDNA"
     gene            55..693
                     /gene="SSI-1"
     CDS             55..693
                     /gene="SSI-1"
                     /codon_start=1
                     /product="stat-induced stat inhibitor-1"
                     /protein_id="BAA21539.1"
                     /db_xref="GI:2274774"
                     /translation="MVARNQVAADNAISPAAEPRRSEPSSSSSSSSSPAAPVRPRCP
                     GVPAPAGDTHFRFRSHSDYRRITRTRSALLDACGFYWGPLSVHGAHERLRAEPVGTF
                     LVRDSRORNCFFALSVKMASGPPSIRVHFQAGRFHLDGNRETFDCLFELLEHYVAAPR
                     RMLGAPLRQRVRPLQELCRQRIVAAVGRENLARIPLNPVLRDYLSSFPFQI"

BASE COUNT      121 a    338 c    274 g    176 t
ORIGIN
```

```
Best Local Similarity 906;  Conservative 99.7%;  Pred. No. 9.4e-154;  Indels 0;  Gaps 0:
Matches 906;  Mismatches 3;
```

```
Qy    781 CTTCTGGCTTGGAGACCTCATCCAACTCTTCAGGGGTGGGGGGTGCTCCCCTCCTGGTGCTC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTTCTGGCTTGGAGACCTCATCCAACTCTTCAGGGGTGGGGGGTGCTCCCCTCCTGGTGCTC 840

Qy   1001 CCTCCGGGTCCCCCGGGGTTGTAGCAGCTTGTGTCTGGGGGCCAGGAGCTTGAATTCCACTC 1060
Db    841 CCTCCGGGTCCCCCGGGGTTGTAGCAGCTTGTGTCTGGGGGCCAGGAGCTTGAATTCCACTC 900

Qy   1061 CTACCTCTCCATGTTACATATTCCCAGTATTCTTTGCACAAACCAGGAGGGGGTGGGGGGT 1120
Db    901 CTACCTCTCCATGTTACATATTCTTTGCACAAATCTTTTGCACAAACCAGGAGGGGGTGGGGGGT 960

Qy   1121 CTCTGGCTTCATTTTTTTCTGCTGTGTGGCAGAATATCCTATTTTTATTTTTTACGCCAGTTTA 1180
Db    961 CTCTGGCTTCATTTTTTTCTGCTGTGCACAAATCCTATTTTTATTTTTTACGCCAGTTTA 1020

Qy   1181 GGTAATAAACTTTATTATATGAAAGTTTTTTTTTTTTAAAA 1216
Db   1021 GGTAATAAACTTTATTATATGAAAGTTTTTTTTTTTTAAAA 1056
```

```
Qy    107 TCAGGCGCAGCCCCGAGCCTCATGGCCGCTATGGCGCCACCGACTGCCCTGCTGGCTCGACTGACTGAGTTAGGAGATGGTA 166
Db      1 TCAGGCGCAGCCCCGAGCCTCATGGCCGCTATGGCGCCACCGACTGCCCTGCTGGCTCGACTGACTGAGTTAGGAGATGGTA 60

Qy    167 GCACGGCCAACCAGGTGGCAGCCGACAAATGCGGAATCTCCCCGGCCGCAGAGAGCCCCGACGGGGCG 226
Db     61 GCACGGCCAACCAGGTGGCAGCCGACAAATGCGGAATCTCCCCGGCCGCAGAGAGCCCCGACGGGCG 120

Qy    227 TCAGAGCCCTCCTCTGGTCCTTCGTCCTGCGCAGGTCAGTTTA 286
Db    121 TCAGAGCCCTCCTCTGGTCCTTCGTCCTGCGCAGGTCAGTTTA 180

Qy    287 TGCCCGGCGGTCCCAGCCCCACCCCCTGGGCGGACACTCACTTCCGGCGACCTTCGGCGTCCCCAC 346
Db    181 TGCCCGGCGGTCCCAGCCCCACCCCCTGGGCGGACACTCACTTCCGGCGACCTTCGGCGTCCCCAC 240

Qy    347 TCCGATTACCCGGCGCCGGCCATCACGCGGACCAGCGGCGCTGCTCCGGACGCCGGCGGCTTCTATTGG 406
Db    241 TCCGATTACCCGGCGCCGGCCATCACGCGGACCAGCGGCGCTGCTCCGGACGCCGGCGGCTTCTATTGG 300

Qy    407 GGACCCCTGAGCGTGCACGGGGCCCACGAGCGGCTGCGTGCCGGAGCCCGTGGCCGGCAACCGC 466
Db    301 GGACCCCTGAGCGTGCACGGGGCCCACGAGCGGCTGCGTGCCGGAGCCCGTGGCCGGCAACCGC 360

Qy    467 TTGGTGCGCGACAGTCGTCAACGGAGAACTGCTTCTTCGCGCTCAGCGTGAAGAATGGCTTCG 526
Db    361 TTGGTGCGCGACAGTCGTCAACGGAGAACTGCTTCTTCGCGCTCAGCGTGAAGAATGGCTTCG 420

Qy    527 GGCCCCACGAGCATCCGGCGCTGCACTTCCAGGCGCAGCGGCTTCCACTTGGACGGAGCGGC 586
Db    421 GGCCCCACGAGCATCCGGCGCTGCACTTCCAGGCGCAGCGGCTTCCACTTGGACGGCAACCGC 480

Qy    587 GAGACCTTCGACTGCCTTTTCGAGCTGCTGGAACACTACGTGGCGGCGCCGCGGCGCGATG 646
Db    481 GAGACCTTCGACTGCCTTTTCGAGCTGCTGGAACACTACGTGGCGGCGCCGCGGCGCATG 540

Qy    647 TTGGGGGGCCCCGGCTGCGGAGCGCGTCGGAGCTGTGTCGCCAGCGC 706
Db    541 TTGGGGGGCCCCGGCTGCGGAGCGCGTCGGAGCTGTGTCGCCAGCGC 600

Qy    707 ATCGTGGCCGCGCGGTGGGATCCCGTGGGTCGCGAGAACCTGGCGCCATCCTCTTAACCGGTACTCCGT 766
Db    601 ATCGTGGCCGCGCGGTGGGATCCCGTGGGTCGCGAGAACCTGGCGCCATCCTCTTAACCGGTACTCCGT 660

Qy    767 GACTACCTGAGTTCCTTCCCCTTCCAGATCTGACGGGCTGCGTGCCGCAGCATTAA 826
Db    661 GACTACCTGAGTTCCTTCCCCTTCCAGATCTGACGGGCTGCGTGCCGCAGCATTAA 720

Qy    827 GTGGGGGGGCGGCCTTATTATTTCTTAATTATTATTAATTATTTTTTCTGGAACCACGTGGG 886
Db    721 GTGGGGGGGCGGCCTTATTATTTCTTAATTATTATTAATTATTTTTTCTGGAACCACGTGGG 780

Qy    887 AGCCCTCCCCGCCTTGGGGTCGGAGGGAGTGGTTGTTGGAGGGTGAGATGCCTCCCACTTCTG 946
Db    781 AGCCCTCCCCGCCTTGGGGTCGGAGGGAGTGGTTGTTGGAGGGTGAGATGCCTCCCACTTCTG 840

Qy    947 GCTGGGAGACCTCATCCCACCTCTCAGGGGTGGGGGGTGTCCCCCTGGGGGGTGGGGGG 1006
Db    841 GCTGGGAGACCTCATCCCACCTCTCAGGGGTGGGGGGTGTCCCCCTGGGGGGTGGGGGG 900

Qy   1007 GGTCCCCCC 1015
Db    901 GGTCCCCCC 909
```

```
RESULT 8

LOCUS       AC002286      12589 bp    DNA             PRI       23-NOV-1999
DEFINITION  Homo sapiens Chromosome 16p13.13 cosmid clone 356d7 from the Los
            Alamos Chromosome 16 specific Supercos cosmid library complete
```

```
ACCESSION   AC002286
VERSION     AC002286.1  GI:2160310
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 12589)
  AUTHORS   Kramer,J.A., Adams,M.D., Singh,G.B., Doggett,N.A. and Krawetz,S.A.
  TITLE     Extended analysis of the region encompassing the PRM1-->PRM2-->TNP2
            domain: genomic organization, evolution and gene identification
            from human chromosome 16p and 16q
  JOURNAL   J. Exp. Zool. 282 (1-2), 245-253 (1998)
  MEDLINE   98390623
REFERENCE   2  (bases 1 to 12589)
  AUTHORS   Loftus,B.J., Kim,U.J., Sneddon,V.P., Kalush,F., Brandon,R.,
            Fuhrmann,J., Mason,T., Crosby,M.L., Barstead,M., Cronin,L.,
            Deslattes Mays,A., Cao,Y., Xu,R.X., Kang,H.L., Mitchell,S.,
            Eichler,E.E., Harris,P.C., Venter,J.C. and Adams,M.D.
  TITLE     Genome duplications and other features in 12 Mb of DNA sequence
            from human chromosome 16p and 16q
  JOURNAL   Genomics 60 (3), 295-308 (1999)
  MEDLINE   99425270
REFERENCE   3  (bases 1 to 12589)
  AUTHORS   Adams,M.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUN-1997) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   4  (bases 1 to 12589)
  AUTHORS   Adams,M.D., Loftus,B.J., Zhou,L., La Bombard,M., Kim,U.J.,
            Venter,J.C., Kramer,J.A. and Doggett,N.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-AUG-1997) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT     Address all correspondence to:
            Mark Adams
            The Institute for Genomic Research
            9712 Medical Center Dr,
            Rockville, MD 20850,
            USA  e-mail address: mdadams@tigr.org. The orientation of the
            sequence is from SP6
            end to T7 end. Genes were identified by a combination of five
            methods including: XGRAIL (available by anonymous ftp from
            arthur.epm.ornl.gov), Genefinder (Phil Green, University of
            Washington), Genscan (Chris Burge),
            http://gnomic.stanford.Edu/~chris/GENSCANW.html) searches of the
            complete sequence against a peptide database, and the EST database
            at TIGR (http://www.tigr.org/tdb/at/at.html). A gene with homology
            to another protein is annotated as the isolog of that protein.
            Genes without peptide homology having spliced EST hits are termed
            'unknown protein'. Genes encoding tRNAs are predicted by
            tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
FEATURES             Location/Qualifiers
     source          1..12589
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /map="16"
                     /clone="Protamine16"
     repeat_region   complement(381..440)
                     /rpt_family="GC_rich"
     repeat_region   complement(751..806)
                     /rpt_family="GC_rich"
     repeat_region   complement(1318..1361)
                     /rpt_family="AT_rich"
     repeat_region   complement(1572..1700)
                     /rpt_family="MIR"
     repeat_region   complement(2627..2730)
                     /rpt_family="MIR"
     repeat_region   complement(2846..2888)
                     /rpt_family="(GAAAA)n"
     repeat_region   complement(3026..3134)
                     /rpt_family="MIR"
     repeat_region   complement(3458..3501)
                     /rpt_family="(CA)n"
     repeat_region   complement(3810..3902)
                     /rpt_family="MIR"
     repeat_region   complement(3936..4240)
                     /rpt_family="AluSx"
     repeat_region   complement(4318..4367)
                     /rpt_family="MIR"
     repeat_region   4394..4455
                     /rpt_family="MIR"
     repeat_region   complement(4569..4864)
                     /rpt_family="AluJo"
     repeat_region   6540..6617
                     /rpt_family="MIR"
     repeat_region   complement(6655..6705)
                     /rpt_family="(CAAA)n"
     repeat_region   6726..7025
                     /rpt_family="AluSx"
     repeat_region   complement(7029..7296)
                     /rpt_family="AluJo"
     repeat_region   complement(7445..7614)
                     /rpt_family="MIR"
     repeat_region   8005..8310
                     /rpt_family="AluSx"
     repeat_region   8311..8358
                     /rpt_family="MIR"
     repeat_region   8359..8659
                     /rpt_family="(GAAA)n"
     repeat_region   8938..9009
                     /rpt_family="AluSx"
     repeat_region   9021..9331
                     /rpt_family="MIR"
     repeat_region   complement(9842..9998)
                     /rpt_family="AluJb"
     repeat_region   10416..10552
                     /rpt_family="MIR"
     repeat_region   10557..10599
                     /rpt_family="MIR"
     repeat_region   10692..10899
                     /rpt_family="POLY_A"
     repeat_region   complement(11234..11542)
                     /rpt_family="AluSx"
     repeat_region   12198..12309
                     /rpt_family="AluY"
     repeat_region   complement(12203..12589)
                     /rpt_family="(CAT)n"
     misc_feature    /note="has high homology to Human nuclear
                     ribonucleoprotein particle (hnRNP) C protein"
                     /evidence=not_experimental
BASE COUNT     3067 a   3233 c   3221 g   3068 t
ORIGIN
```

```
Query Match             58.8%;  Score 725.6;  DB 40;  Length 12589;
Best Local Similarity   80.7%;
Matches 1004;  Conservative   0;  Mismatches 204;  Indels  36;  Gaps  12;
                        Pred. No. 1.3e-121;

Qy      6 CTCAAGCTCCGGCGCGGGAATTCTGCGTGCCGGTCCTTGGGGTCTGTTGCCGGGCCT    65
            | | ||  |||||  ||
Db    131 CCCTAGGGACGGGCGGGGTTCCGCCGGGAGCGCGGCGCCCTTCCGGAGTTGTTCCGGGCCGGGCACACTT  190

Qy     66 GTGCCACCCGGACGGCCGGGCTCACTGCCTCTGTCTC-CCCCACACAGCCAGCCCCGGGACG  124
            | |||||  | || |||| | |  |||| ||||  | | |  |  | | | ||| ||||
Db    191 GCCCGGGCGCCCGCAGCGCCCCCAGCTCACCGCTTTGTCTCCGCGGCCAGCCGACCCCCGGGACG  250

Qy    125 CTATGGGCCCACCCCCTTCCAGCTGGCCCCCTCGGAGTAGGATAGGATGGTAGCACGCCAAACCAGGTGGCA  184
            |||||||||||||||||| | | || ||  |||| || |||||| |||||| ||||||| |||||
Db    251 CTATGGGCCCACCCCCTCCGGGCTGGCCCCTTCTCTGTPAGGAGTGGTAGCACACACAACCAGGTGGCA  310

Qy    185 GCCGACAATGCGAATCTCCCCCGGCCAGCAGAGCCCCGACGGCGGTCAGAGACCCCTCCTCGTCC  244
            ||||||||| | || |||||| | |||||||||||||||||| ||| |||
Db    311 GCCGACAATGCAGTCTCCACACAGCAGCAGAGCACCCCGACGGGCGGCCAGAACCTTCC---TCC  367

Qy    245 TCGTCTGTCGTCTCGCCAGGCGGCCCCCGTGGCGTCCCGGGCCCCGGGGTCCCAGCC  304
```

```
RESULT  9
LOCUS       HSAC002047
DEFINITION  Homo sapiens clone Protamine16. *** SEQUENCING IN PROGRESS ***, 2
            unordered pieces.
```

```
ACCESSION   AC002047
VERSION     AC002047.1  GI:2062143
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 25417)
  AUTHORS   Adams,M.D., Loftus,B.J., Zhou,L., Phillips,C., Brandon,R.C.,
            Fuhrmann,J., Kim,U.J., Kerlavage,A.R. and Venter,J.C.
  TITLE     Human chromosome 16p13 BAC clone CIT987SK-Protamine16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 25417)
  AUTHORS   Adams,M.D. and Loftus,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-1997) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
REFERENCE   3  (bases 1 to 25417)
  AUTHORS   Adams,M.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-1997) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT     NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *     1       12828: contig of 12828 bp in length
            *     12829          gap of unknown length
            *     12829  25417: contig of 12589 bp in length.
FEATURES             Location/Qualifiers
     source          1..25417
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Protamine16"
BASE COUNT     6160 a   6571 c   6525 g   6159 t   2 others
ORIGIN
```

```
Query Match             58.8%;  Score 725.6;  DB 32;  Length 25417;
Best Local Similarity   80.7%;  Pred. No. 1.2e-121;
Matches 1004;  Conservative  0;  Mismatches 204;  Indels 36;  Gaps 12;
```

```
Qy    368  TCTTCCTCTCCTCTGCCCGCGCGGGCGCGCGGC  427
Db    305  CCAGCGCCCTGCGCGACACGCACTTCACTTCCGCCGACCACTTCCGGTCCGGTGCCCGCGGGTCCCGCGGGCGCGGTGCCCGCGGGTCCGGCGCATC  364

Qy    428  CCGGCCCCGCGGCGACACGCACTTCCCGCACCATTCCGTTCGCACGCCCGATTACCGGCCATC  487
Db    365  ...  424

Qy    488  ACGCGCGCACCAGCGCCTCCTGGACGCCGGTGCAC  547
Db    425  GGGGGCCACGAGCGCGTGCGCCGGCGCCGGTGCCCGACAGTCGT  484

Qy    548  GGGGGCCACGAGCGCGTGCGGCGGCGCCGGTGCCCGACAGTCGT  607
Db    485  CAACGGGAACTGCTTCTTCGCCTCACGTGTGAGGATGGCTTGGGCGCCCACGAGCGCCGC  544

Qy    608  CAGCGGCGAACTGCTTTTTCGCCCTTACGGTGAAGATGCCCTGGGACCCACCACGAGCCGC  667
Db    545  GTGCGCACTTTCAGGCCGCGGTTTCACCTGGATGGCAGCCGCGAGAGCTTGGACTGCCTC  604

Qy    668  GTGCGCACTTTCAGGCCGCGGTTTCACCTGGATGGCAGCCGCGAGAGCTTGGACTGCCTC  727
Db    605  TTCGAGCTGCTGGAGCACTAGGTGGGCGGCGCGCGCCGCGCANGTTGGGGCGCCCGCGCGC  664

Qy    728  TTCGAGCTGCTGGAGCACTAGGTGGGCGGCGCGCGCCGCGCATGCTGGGGGCGCCCGCTGCGC  787
Db    665  CAGCGCGCCGCGTGCCGCGCGCTCAGGAGCTGTGTCGCCAGCGCCATCGTGGCCGCCCGTGGGT  724

Qy    788  CAGCGCGCCGCGTGCCGCGCGCTCAGGAGCTGTGTCGCCAGCGCCATCGTGGCCGCCCGTGGGC  847
Db    725  CGCGAGGAACCTGGCGCCGCCGCATCCCTTAAACCCGGTACTCCGTGACTACCTGGAGTTCCTTC  784

Qy    848  CGCGAGAAACCTGGCGCTGCGCATCCCCCGTCACCCCGTCCTCCGGCGCTGGACGCTGCCTGG  907
Db    785  CCCTTCCAGATCTGACCGGCGTGCCGCACGCATTAAGTGGGGGCGGCGCGGCCCTTATTAT  844

Qy    908  CCCTTCCAGATTGACGGGCAGGGCGCGCCCGTGTCACGCAGCAGCATTAAACTGGGGATGCCCGTGT  967
Db    845  TTCTTATTATTAATTAATTATTATTTTTCTGAAACCACGTGGGAGCCCCTCCCC-GCCTGGG  903

Qy    968  TATT------TTCTTATTACTTGCCTGAAACCATGTGGGTACCCTCCGGCCTGGGAGA------  1018
Db    904  TCGGAGGGGAGTGGGTTGTG---GAGGGTGAGATGCCTCCCACTTCTCGGCCTGGAGA-----C  955

Qy   1019  TTGGGAGGGAGCGGGATGGGTGTTGTAGGGGCGGCGGCCGCCGCGCGTCCCGGGCGGGGGTCGTGTTGGGGCCGCGGCC  1078
Db    956  CTCATCCCACCTCTCAGGGGTGGGGGTGCT----CCCCTCCTGGTCGTCCCCTCCCTCCCGGGGTCC  1011

Qy   1079  CGCAGAGACCCTTCTCACCCTCTTGAGGGGGGTCCTCCCCATCAGGGGCAGGCCACT  1138
Db   1012  CCC---CTGGTTGTGAGCAGCT----TCTGTCTGGGGCCCAGGAGGAATT---CCACTCCTA  1063

Qy   1139  CCCTGGTTGTTGTGTAGCAGCTTAACTGTATCTTTTGGGAGGCCAGGACCTGAAACTCGGCCACCTCCTA  1198
Db   1064  CCCTCTCCCATGTTACATATCCCACAGTATCTTTTGCACAAACCAGGGG-TCGGGGAGGGTCT  1122

Qy   1199  CCTCTTCATGTTTACATATACCCAGTATCTTTTGCACAAACCAGGGGTTGGGGAGGGGTCT  1258
Db   1123  CTGGGCTTCAATTTTCTGCTGTGCAGAGATATCCTATTTTATATTTTTACAGCCAGTTAGG  1182

Qy   1259  CTGGGCTTATTTTCTGCTGTGCAGAATCCTATTTTTATAT-TTTTAAAGTCAGTTTAGG  1317
Db   1183  TAATAAACTTTTATTATGAAAAGTTTTTTTTTTTTTTAAAAAAATAA  1226

Qy   1318  TAATAAACTTTATTATGAAAAGTTTTTTTTTTTTTTAAAAAAAATAA  1361
```

```
Qy      6  CTCAAGCTCCGGGCGGATTCTGCGTCCGCGTCCGCTCTCGCGTCTCCGGTCTGTTGGGGTCGGCGGGCCCT  65
Db    370  CCCTAGGGCCCGGGTCCCGCGGGGAGCGCGGCCCGGCCGGGTGTTCCGGCCGGGCCGGCACACT  429

Qy     66  GTGCCACCCGGACGCCCGGGCTCACTGCCTCGTCTC--CCCCATCAGGCGCAGCCCCGGACG  124
Db    430  GCCCGGCCCCCCAGCCGCCCCAGCTCACCCGCGTTTGTCTCCTCCCGCGCAGCGCACCCCGGGACG  489

Qy    125  CTATGGGCCCACCCCTCCGGCGCTGGCCCCTCGACTAGGAGGATGGTGAGCACGCAACCAGGTGGCA  184
Db    490  CTATGGGCCCACCCCTCCGGCTGGCCCCTTCGTGAGGATGGTAGCACACAACCAGGTGGCA  549

Qy    185  GCCGACAAATGGGATCTCCCGCGCAGCAGAGAGCCCCGACGGCGGTCAGAGCCCCTCGGTCC  244
Db    550  GCCGACAATGCAGTCTCCACACAGCAGAGGCCCCGACGGCCAGAACCTTCC---TCC  606

Qy    245  TCGTCTTCGTCCTCCGCCAGGGGCGCCCCGGCGCCCCGTGGGCCCCGGACCCCGGCC  304
Db    607  TCTTCCTCTCCTCTGCCCGCCGCCCCGTGGGCCCCGGACCCCGGCC  666

Qy    305  CCAGCCCCTGCGGCGACACTCACTTCCCGCACCACTTCCGGTCCGGTGCCCGCGGGTCCCGCGGGCGCGGTGCCCGCGGGTCCCGGCCATC  364
Db    667  CCGGCCCCGCGGCGACACGCACTTCCCGCACCATTCCGTTCGCACGCCCGATTACCGGCCATC  726

Qy    365  ACGCGCGACCACGCGCCTCCTGGACGCCGCTGGGGCTTCTATTTGGGGGACCCCTGAGCGGTGCAC  424
Db    727  ACGCGCGCACCAGCGCGCCTCCTGGACGCCCTGCGGGATTCTACTGGGGCGTGCAC  786
```

RESULT 10
LOCUS   AC007220   175950 bp   DNA   HTG   06-APR-1999
DEFINITION Homo sapiens chromosome 16 clone 396B14, *** SEQUENCING IN PROGRESS
            ***, 2 ordered pieces.
ACCESSION   AC007220
VERSION   AC007220.1 GI:4567168
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 175950)
  AUTHORS   Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
            Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
            Bryant,J., Tesmer,J., Meincke,L., Longmire,J., White,S., Tatum,O.,
            Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
            Han,C. and Deaven,L.
  TITLE   Sequencing of Human Chromosome 16p13.3
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 175950)
  AUTHORS   Ricke,D.O.
  TITLE   Large Scale Sequence Analysis and Annotation with the Sequence
            Comparison Analysis (SCAN) System
  JOURNAL   Unpublished
REFERENCE   3 (bases 1 to 175950)
  AUTHORS   Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
            Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
            Bryant,J., Tesmer,J., Meincke,L., Longmire,J., White,S., Tatum,O.,
            Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
            McMurry,K., Han,C. and Deaven,L.
  TITLE   Direct Submission
  JOURNAL   Submitted (06-APR-1999) Center for Human Genome Studies, DOE Joint
            Genome Institute, Los Alamos National Laboratory, MS M888, Los
            Alamos, NM 87545, USA
COMMENT   * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *   1 29280: contig of 29280 bp in length
            *   gap of unknown length
            * 29281 175950: contig of 146670 bp in length.
FEATURES   Location/Qualifiers
     source   1..175950
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="396B14"
                     /chromosome="16"
BASE COUNT   44159 a   45076 c   42509 g   44199 t   7 others
ORIGIN

Query Match   58.6%; Score 723.6; DB 33; Length 175950;
Best Local Similarity   80.7%; Pred. No. 2.7e-121;
Matches 1002; Conservative   0; Mismatches 204; Indels 36; Gaps 12;

Qy   425 GGGGCGCCACGAGCGGCGTCGTGCCGACCGGCTGGCGACCTTCTTGGTGGTGCGGCGACAGTCGT 484
          ||||||||| |||||||| ||| ||||| || |||||||||| || |||||| |||||| |||
Db   787 GGGGCGCCAGAGCGGCGCTGGCGGCGGCTGGGGCACCTTCCTGGTGGCGGGGACGAGCACAGCGC 846

Qy   485 CAACGGAACTGCTTCTTCTCGCCCTCACGTGAAGATGGCTTGGGCGCCCACAGCATCCGC 544
          |||||| |||||||||||||| |||| |||| |||||||| |||||| |||| |||||||
Db   847 CAGCGGAACTGCTGCTTTTCGCCCTTAGGTGATGATGGCCTGGGACGAGCACGAGCATCGC 906

Qy   545 GTGCACTTCAGGCGCGCCGCTTCCAGCTTGGACGGGCGGGAGACTTCGACTGCCTT 604
          ||||||| |||| || |||| ||| ||| |||| ||||||| |||| |||| |||
Db   907 GTGCACTTTCAGGCGCGCTTTCACCTGATGGCAGCGCCGCGGAGCCTTCGACTGCCTC 966

Qy   605 TTCGAGCTGCTGGAGCACTACGTTGGCGGGCGGGAGGCGCCGCGCCCGGGGCGTCCGC 664
          ||||| || || ||| || |||||||| |||||||||| ||| ||||||||||||||
Db   967 TTCGAGCTGCTGGAGGACTACGGTGGCGGGCGGGAGGCGCCGCGCGCCGGGGCGTCCGC 1026

Qy   665 CAGCGCGCCGCGCGCCGGGGCCGGCGCGGGGCTCAGGAGACTTCGCGCCGCGTGGGT 724
          |||||| |||||| |||||||||||||||||| ||||||| ||| ||||||| |||
Db   1027 CAGCGCGCCGCGCTGCGGGGACCGGCGCGGGGCCCAGGAGACTGGCGCCGCGTGGGC 1086

Qy   725 CGGCGAGAACCTGGGCGGCGCCATCCCTCTTAACCGGTACTACCTGAGTTCCTTC 784
          || ||| ||||| |||||||||| |||| ||||||||| |||| |||| |||||
Db   1087 CGCGGAGAACCTGCGCGGCGCCATCCCCTCTTCACCCGGTCCTCGCTGAGTTCCTTC 1146

Qy   785 CCCTTCCAGATCTGACCGGGCGTGTGCGCCAGCATTAAGTGGGGGGCGGGCCTTATTAT 844
          |||||||||||| |||| ||| |||||||||||| || ||||||||| ||| ||| ||
Db   1147 CCCTTCCAGATTGACCGGGCGTGCTGCGCCAGCAGTAACTGGATGCCGGTGCCTGGCTGT 1206

Qy   845 TTCTTATTATTAATTATTATTTTTCTGAACACGTTCGAGTTCCTTC 903
          || |||||||| ||||||||| ||| || |||||| ||||||||||
Db   1207 TATT------TTGTTATTACTTGCCTGGAACCATGTGGGTACCCTCCTTC 1257

Qy   904 TCGGAGGAGGAGGTTGTG---GAGGGTGAGATGCCTCCCACTTCTCTGGCTGAGA-----C 955
          ||||||||||||||| || ||||||||| |||||| ||||| ||||||||||| |
Db   1258 TTGGAGGAGGCGGATGGGGTGTGAGGGCGGCCCCTCCCTGCCTGAGACCAGGC 1317

Qy   956 CTCATCCACCTCTCAGGGTGNGGGGTGCT-----CCCCTCCTGGTGCTCCCTGGGGGTCC 1011
          ||||||||||| |||||| | ||| |||| |||| ||||| |||| ||||||||||
Db   1318 CGCAGACCCTTCTCACCTCTTGAGGGGTCCTCCCCCCGTGGTGCTCCCCTCTGGGTCC 1377

Qy   1012 CCC--CTGGTTGTTGTAGCAGCT-----TGTGTCTGGGGGCCAGGACCTGAATT--CCACTCCCTA 1063
          ||| | ||| || || ||| |||||||||||||| ||||| || || |||||||||
Db   1378 CCCTGGTTGTTGGTAGCAGCTTAACTGTATCTGGACGAGCCTGAACTCGGCACCTCCTA 1437

Qy   1064 CCTCTCCATGTTTACATATTCCAGGTATCTTTGCACAAACCAGGGG-TCGGGAGGGGTCT 1122
          |||||||||||| |||||| |||||||| |||||| ||||||||| ||||||||||||||
Db   1438 CCTCTTCATGTTTACATATACCAGGTATCTTTGCACAAACCAGGGGGTGGGGAGGGTCT 1497

Qy   1123 CTGGCTTCATTTTCTGCTGTGTCAGAAATATCCTATTTTATATTTTACAGCCGGTTTAGG 1182
          |||||||| |||||||| ||||||||||| ||| ||||| ||| ||||||| ||||||||
Db   1498 CTGGCTTATTTTTCTGTGGCAGAAATCCTATTTTATAT-TTTTTAAAGTCAGTTTAGG 1556

Qy   1183 TAATAAAACTTTATTTATTATGAAAAGTTTTTTTTTTTTTTAAAAGAAAAATAA 1226
          ||||||||| ||||| ||| ||||||| |||||||||||||||| |||| |||||
Db   1557 TAATAAAACTTTATTTATGAAAGTTTTTTTTTTTTTTAAAAAAATAA 1600

Qy   6 CTCAAGCTCCGGGCGGGATTCTGCGTGCGCTCGCCGCTTGGGGTTGTTGGCGGGCT 65
          |||||||||| ||| || || ||| ||||| |||| | || ||||| |||
Db   5707 CCCTAGGCGGGGTCGGGCCGGGGAGCGGGCGCCGCCCCCCCCGGAGGGGCACACT 5766

Qy   66 GTGCCACCGGACGCGCCCGGGTCACTGCCTCGTGCTC-CCCATCAGCGGCCCCGGACG 124
          |||||||||||||| |||| | |||||| || | || |||| ||| |||| |||||
Db   5767 GCCCGGCCCCCGCAGCGCCCAGCTCACCGGCTTGTCTCTCCCCACCGCCACCCCCGACG 5826

Qy   125 CTATGGGCCACCCCTCCAGCCTGGCCCCCTCGAGTAGGATGGTAGCACCGGCA 184
          ||||||||| ||||| |||| |||| |||||||| |||| |||||| ||||||
Db   5827 CTATGGCCCACCCCCTCGGCGTGGCCCCTTCTCGTAGGATGGTAGCACAACCAGGTGGCA 5886

Qy   185 GCCGACAATGCGATCTCCCCGCGCAGACAGACCCGACGGGGGTCAGAGCCCTTCTCGTCC 244
          |||||||||||| ||| |||||||| |||| ||| |||| ||||||||| ||| |||||
Db   5887 GCCGACAATGCAGTCTCCACAGCAGAGAGCCCCGACGGGGGCCCAGAACCTTCC--TCC 5943

Qy   245 TCGTCTCTTCGTCTCTCGCCAGCCGCCCCCCCGGGTCCCCCCGCCTCCTCCCCAGCC 304
          ||| ||||| ||||||| |||| |||||| |||| ||||||||| |||| ||||||
Db   5944 TCTTCCTCCTCCTGCGCCGCGGGGCCCCCGCCGGGGTCCCGGGCGGTCCGCCAGCC 6003

Qy   305 CCAGCCCCCTGGGGACACTCACTTCCCGACCTTCCCCACTCCGATTACCGGCCATC 364
          |||||| || ||||| |||||| | ||||| |||||||| ||||| ||||| ||||
Db   6004 CCGGCCCCCGCAGCGGACACGCACTTCCCGACCTTCGTCCACTCCGATTTGCTGCCATC 6063

Qy   365 ACGCGGACCAGGCGCGCTCCTGGACGCCTGCGGACCCCTGACGGTGCAC 424
          ||||||||| | | ||||| |||||||| ||||||||||||||||||||
Db   6064 AGGCGCGCCCAGGCGTCCTGGACGCCTGCGGATTCACTGGGGACCCCTGAGCGTGCAC 6123

```
RESULT 11

LOCUS       AB000734    1216 bp    mRNA           PRI    18-NOV-1997
DEFINITION  Homo sapiens mRNA for TIP3, complete cds.
ACCESSION   AB000734
VERSION     AB000734.1  GI:2627028
KEYWORDS    tip3; TIP3.
SOURCE      Homo sapiens blood B-lymphocytes cDNA to mRNA.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1216)
AUTHORS     Mano,H.
TITLE       Direct Submission
JOURNAL     Submitted (27-JAN-1997) to the DDBJ/EMBL/GenBank databases.

REFERENCE   2 (sites)
AUTHORS     Ohya,Ki., Kajigaya,S., Yamashita,Y., Miyazato,A., Hatake,K.,
            Miura,Y., Ikeda,U., Shinada,K., Ozawa,K. and Mano,H.
TITLE       SOCS-1/JAB/SSI-1 can bind to and suppress Tec protein-tyrosine
            kinase
JOURNAL     J. Biol. Chem. 272 (43), 27178-27182 (1997)
MEDLINE     98001695

FEATURES             Location/Qualifiers
     source          1..1216
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="B-lymphocytes"
                     /tissue_type="blood"
     gene            155..790
                     /gene="tip3"
     CDS             155..790
                     /gene="tip3"
                     /note="similar to U88325.SOCS-1, AB000676:JAB,
                     AB000710:SSI-1"
                     /codon_start=1
                     /product="TIP3"
                     /protein_id="BAA23521.1"
                     /db_xref="GI:2627029"
                     /translation="MYAHNQVAADNAVSTAAEPRRPEPSSSSSSSPAAPARPRPCPA
                     VPAPAPGDTHFRTFRSHADYRRITRASALLDACGPYWGPLSVHGAHERLRAEPVGTFL
                     VRDSRQRNCFFALSVKMASGPTSIRVHFQAGRFHLDGSRSSFDCLFELLEHYVAAPRR
                     MLGAPLRQRVRPLQELCRORIVATVGRENLARIPLNPVLRDYLSSFPFQI"
     polyA_site      1216
                     /note="12 A nucleotides"
BASE COUNT     172 a    449 c    355 g    240 t
ORIGIN

Query Match            58.28;  Score 718.6;  DB 9;  Length 1216;
Best Local Similarity  81.1%;
Matches  992;  Conservative  0;  Mismatches 194;  Indels 37;  Gaps 12;
```

```
Qy    425 GGGGGCCACGAGCGGCGGTGCGTGCCGGAGCCCGTGGGGCCGGCCACCTTCTTGGTGGCGGCAGTGT  484
          ||||||| |||||||||| ||||||||| |||| ||||||||||||| || |||||||||||| ||
Db   6124 GGGGGCACGAGCGGCGGCTGCGTGCCGAGCGCGTGGGGCCACCTTCCTGGTGCCGCAGCAGCCGC   6183

Qy    485 CAACGGAACGCTTCTTCCGGCCTCAGCGTGAAGATGGCTTCGGGCCCCCAGAGACATCCGC       544
          || |||||||  || ||  |||||||||| ||||||| |||| ||| || ||||| ||||
Db   6184 CAGCGGAACTGCTTTTTCGCCCTAGCGTGAAGATGGCCTCGGGACCCCACGAGCATCCGC       6243

Qy    545 GTGCACTTCCAGGCCGCCGGCCGCTTCACTTGGATGGCAGCTCGGACTCGCCTT           604
          |||||||||||| || |||| || || |||| ||  || |||||| |||||||
Db   6244 GTGCACTTCCAGGCCGCCGGCCGCTTCACCTGGATGGCAGCCCGGACTCGACTGCCTC       6303

Qy    605 TTCGCAGCTGCTGGAGCACTACGGTGGCCGCGCGGGCGGCCCGGCCCCGGCTGCGC       664
          |||||||||||||||||| ||||| ||||||||||||| || |||||||||||||
Db   6304 TTCGCAGCTGCTGGAGCACTACGGTGGCCGCGCGGGCGGCCCGGCCCCGGCTGCGC       6363

Qy    665 CAGCGCCGCGGCGCCGCGGCCGCGGGCGGCGGCGCGGGCGGCGGCGGCCCCGGGT       724
          ||||||||||| ||||| ||||||||||| ||||||||||||||||||| |||||
Db   6364 CAGCGCCGCGGCGCTGCGGCCGCGGGCGGCCGCGCGGGCGGCGGCGGCCCCGGGGC     6423

Qy    725 CGGCGAGAACCTGGCGCGGGCCGCATCCTTCTTAACCCGGTACTCGAGTTCCTTC       784
          |||||| ||| || || ||| || || |||| |||||||| ||| || |||||
Db   6424 CGGCGAGAAACTGGCGCGCATCCCCCCTCAACCCGGTACTCCGCCGAGTTCCTTC       6483

Qy    785 TTCTTATTATTAATTAATTATTTTCTGGAAACCACGTGGAGGCGGCCTTATTAT       844
          | |||| || || |  ||| ||  |||||||||| ||| || ||||||||||||
Db   6484 CCCTTCCAGATTGACGGCGAGCGCCGTGCACGCAGCATTAACTGGGGGCCCTTC      6543

Qy    845 TTCTTATTATTAATTAATTATTTTTCTGGAAACCACGTTCCCCC--GCCTGGG       903
          | |||||||||| |  ||||||||| |||| |||||| |||||   |||||||
Db   6544 TTGTTATTACTTGCCTGGAACCATGTGGGTACCCCTCCCCCGCGCCCGGGCCTGGG  6594

Qy    904 TCGGAGGAGGAGTGGTTGTG---GAGGGTGAGATGCCCCACTTCTCTGGCTGGAGA-----C   955
          |||| || || || ||||||   ||||| || ||  ||||| ||||||||||||||     |
Db   6595 TTGGAGGAGGAGTGGTTAGGGGGTGAGGGCGGCGAGCCCCTTCTCGGCTGGAGGAGGC      6654

Qy    956 CTCATCCCACCTCTCAGGGGTGGGGGTGGCT----CCCCTCCGTGTGCTCCCCTCCGCGGGTCC 1011
          |  ||||| ||| | |||||||||||||||    ||||||||||||||||||||||||||||
Db   6655 CGCACACCCCTTCTCACACCTCTTGGAGGGGGTCTCCCCTCCGTGTGCTCCCCTCTGGGTCC   6714

Qy   1012 CCC---CTGGTTGTAGCAGCT------TGTGTCTGGGGGGGCCAGGAGACCTGAATT--CCACTCCTA 1063
          |||   |||||||||||||||      |||| |||||||||||||||||||||||   ||||||||
Db   6715 CCCGGTTGTTGTAGCAGCTTTAACTGTATCTGGGGGGCCAGGAACCTCGGAACTCGACTCGACTCCTA 6774

Qy   1064 CCTCTCCATGTTACATATTCCCAGTATCTTTGCACAAAACCAGGGG-TCGGGGAGGGTCT      1122
          ||||| || || || || ||||| |||||||||||| || |||||  ||| |||||||||
Db   6775 CCTCTTCATGTTACATATCCCCAGTATCTTTGCACAAAACCAGGGGTTGGGGAGGGTCT      6834

Qy   1123 CTGGCCTTCATTTTTCTGCGTGTGCAGAGAATATCCTATTTATATATTTATTTACAGGAGTTTAGG 1182
          |||||||| | |||| |  ||| |||||| | |||| ||||| |  |||| || |||| ||||||
Db   6835 CTGGCCTTATTTTTTCTGCTGTGCAGAATCCTATTTATTTT-TTTTTTAAAGTTCAGTTTAGG    6893

Qy   1183 TAATAAACTTTATTATATGAAAAGTTTTTTTTTTTTTTTAAAAGAAAATAA              1224
          || ||||| || |||| ||| || ||| ||||||||||| || ||||||||
Db   6894 TAATAAACTTTATTTGTTGCTATGGAAATCTATTTTTATAT-TTTTTAAAGTTCAGTTTAGG     6935
```

```
Qy     14 CCGGGGGGGATTCTGCGTGTGCCGCCTCTCGCCTCCTTGGGGGGTCTGTTGGGCCGGGAGGCCACC   73
          |||| | | | ||| || |||||||||| |||||||||||||||| ||||| || |||||||||
Db      6 CTGCACGGCGTCCTGGCCCCCCGGAGAGCATGCCGAGAGCGCCCGGCGCGCCCCGGGAGCGCC     65

Qy     74 CGGACGCCGCGGCTCACTGCCCTCTGTCTCCCC--CCATCAGCGGCAGCCCCGGCAGCGCGGCTATGGC 131
          ||||| || |||| ||||| | |||||||||||   ||||||| ||||||||||||| || ||||||||
Db     66 CCGCCGTCCCGGCGCGCCCCGGGCGGGTCCCCGGGCCCACGCGCCCAGCCCCCACGGCTATGGC      125

Qy    132 CCACCCCTCCCAGCTGCCCCCTCGAGTAGGAGGATGGGTAGCACGCCAAACCAGGTGGCAGACA     191
          |||||||| || ||||||||| || |||||||||||||| || || |||||| || ||||||||
Db    126 CCACCCCTCCCCGGCTGCCCCCTTCTGTAGGGATTAGGAGGATGGTAGGCACGACAACCAGGTGGCAGACA 185

Qy    192 ATGCGATCTCCCCGGCGCAGAGCCCCGAGGCGGTCAGAGCCCCTCCTGTCCCTGGTCTT       251
          ||| || | || ||  || |||| || |||| ||||| || ||||| || |||| |||
Db    186 ATGCAGTCTCCACAGCAGCAGAGCCCCGGGGGCGGTCAGGAGCGCCCCTCTTCTGCAGCCGACA   242

Qy    252 CGTCCTGCCACGGCCCCCGTGCGTCCCCGGCCCGTGCCCGGGTCCAGCCCCAGCCC        311
          |||||||| | ||||||| |||||||| |||||| || |||| | |||||||||||
Db    243 CCTCCTCCGCGCCGGCGCGTGCGTCCCCGGCCCGTGCCCGGGTCCCGCCCCAGCCC       302

Qy    312 CTGGCGCACTCACTTCCGCGCACCTTCCCGCTCCCGCGCCGCCCGGCCGATCACGCACGCGGA 371
          ||||| || |  |||||| | || |||||||||| || ||||||| ||| || || || ||
Db    303 CCGGCGCACGCGCTTCCGCACATTCCGTTCGCACCCCGGATTACCCGCCGCATCACGCACGGCG  362

Qy    372 CCAGCCGCGCCGCTCCTGGACGGCTCGGGGTTCTCATTGGGGGACCCTGACCGTGGCCGGGGC 431
          ||||| ||| || || ||| ||| ||||| | ||| || ||||||||||||||||||||||||
Db    363 CCAGCGCGCCTCCTGGACGCCCTGCCCTGGATTCTCACTGGGATTCACTGGGGGCCCTGGGGC     422

Qy    432 ACGAGCGGGCTCCGTGCCGAGCCCGTCGGGCCACCTTCCGGACGCGCCACGGGGC         491
          ||||| ||| ||| | |||| || || |||||| |||||| || ||| |||||||
Db    423 ACGAGCGGGCTCGCGGCGGGCGGCAGCCGGGCCACTTCCGGCGGCGCCGACAGCGGA       482
```

```
AUTHORS   Starr,R., Willson,T.A., Viney,E.M., Murray,L.J.L., Rayner,J.R.,
          Jenkins,B.J., Gonda,T.J., Alexander,W.S., Metcalf,D., Nicola,N.A.
          and Hilton,D.J.
TITLE     Direct Submission
JOURNAL   Submitted (04-FEB-1997) Cancer Research Unit, The Walter and Eliza
          Hall Institute of Medical Research, PO Royal Melbourne Hospital,
          VIC 3050, Australia
FEATURES            Location/Qualifiers
     source          1..1094
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="B lymphocyte"
                     /cell_line="Raji cell line"
     gene            24..659
                     /gene="SOCS-1"
     CDS             24..659
                     /gene="SOCS-1"
                     /codon_start=1
                     /product="suppressor of cytokine signalling-1"
                     /protein_id="AAB62401.1"
                     /db_xref="GI:2245384"
                     /translation="MVAHNQVAADNAVSTAAEPRRRREPSSSSSSSPAAPARPRPCPA
                     VPAPAPGDTHFRTFRSHADYRRITRASALLDACGFYWGPLSVHGAHERLRAEPVGTFL
                     VRDSRQRNCFFALSVKMASGPTSIRVHFQAQRFHLDGSRESFDCLFELLEHYVAAPRR
                     MLGAPLRQRRVRPLQELCRQRIVATVGRENLARIPLNPVLRDYLSSSFPFQI"
BASE COUNT      167 a    381 c    313 g    233 t
ORIGIN
```

```
RESULT 12
LOCUS       HSU88326      1094 bp    mRNA            PRI       06-JUL-1997
DEFINITION  Human suppressor of cytokine signalling-1 (SOCS-1) mRNA, complete
            cds.
ACCESSION   U88326
VERSION     U88326.1  GI:2245383
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1094)
  AUTHORS   Starr,R., Willson,T.A., Viney,E.M., Murray,L.J.L., Rayner,J.R.,
            Jenkins,B.J., Gonda,T.J., Alexander,W.S., Metcalf,D., Nicola,N.A.
            and Hilton,D.J.
  TITLE     A family of cytokine-inducible inhibitors of signalling
  JOURNAL   Nature 387 (6636), 917-921 (1997)
  MEDLINE   97345633
REFERENCE   2  (bases 1 to 1094)
```

```
Query Match      55.9%;  Score 690.2;  DB 10;  Length 1094;
Best Local Similarity   83.5%;          Pred. No. 3.2e-115;
Matches  924;  Conservative   0;  Mismatches  148;  Indels  35;  Gaps  11;
```

```
Qy  138 CTCCAGCTGGCCCCCTCGAGTAGGATGGTAGCACGCAACCAGGTGGCAGCCGACAATGGA 197
        ||||   |||| |||| |||| ||||||| ||| ||||| || || || |||||||| |||
Db    1 CTCCGGCTGGCCCCTTCTGTAGGATGGTAGCACACAACCAGGTGGCAGCCGACAATGGAG 60

Qy  198 TCTCCCCGCCAGCAGAGCGCCCGACAGGCCGGTCAGACCCCTCGTCCGTCTTCGTCTCCT 257
        |||||| |||||||||| || ||| ||| || |||| ||||  |||||||||||||||||
Db   61 TCTCCACAGCAGCAGAGCCCCGAGCCCCAGAACCTTCC----TCCTCTTCCCTCCT 117

Qy  258 CGCCAGGGGCCCCGTCGCTCCCCGGGCCCTGCCCGGGGGTCCCCAGCCCTGCG 317
        ||||  |||| ||||| ||| ||| ||||||| ||||| || |||||| ||||
Db  118 CGGCCCGCGGCCCCCGTCGCCCGGTCCCCGGCGTCCGGGGTCTCGGCTT 177

Qy  318 ACACTCACTTCCGCACCTTCCCGCTCCCACTCCGATTACCGGCCGCCATCGACCACC 377
        ||| |||| ||| ||| || ||| || ||||  |||||| |||| || |||| |||
Db  178 ACACGCACTTCCGCACCATCCGTTCGCACGCGCGGATTACCGGGCGCCATCACCAGCG 237

Qy  378 CGCTCCTTGGACGGCCTGCGGCGCTTCTATTGGGAGGACCCTGAGCGTGCAGGTGCACGG 437
        ||||| || || || ||||| ||||||| ||||||| ||||| |||| || |||||||
Db  238 CGCTCCTTGGACGGCCTGCGGGGGGATTCTACTGGGGGCCCGGAGCGTGCACGGGGCCACGGAG 297

Qy  438 GCGTCGGCTGTGCCCAGCCCGCCACCTTCTTGGTGCGCGACAGTGCGTCAACCGAACTGCT 497
        |||| || ||| ||| ||| || ||||| || |||||| |||| || || ||||| |||
Db  298 GGCTGCCGCGGAGCGCCGCCTTCCTGGTGCGCGACAGCGCCCAGCGCAACTGCT 357

Qy  498 TCTTCCGGCGCTCAGGCGTGAAAGATGGCTTCGGGGCCCCGAGCGAGCATCCGGGTGCACTTCCAGG 557
        | ||| ||||||| ||| || || |||| ||||| |||| ||||||| |||| || |||  ||||
Db  358 TTTTCGCCCTTAGCGTGAAAGATGGCTTCGGGACCACCAGAGCATCCGGGTGCACTTCAG 417

Qy  558 CCGGCCGCCGCTTCCACTTGGACGGCCAGCGCGAGACCTTCGACTGCCTTTTCGAGCTGCTGG 617
        ||||| || ||||| || || || |||| ||| |||| ||| ||||||||||| || |||||
Db  418 CCGGCCGCCGCTTCCACTTGGACGGCAGCGGGAGAGCTTCGACTGCCTCTTCGAGCTGCTGG 477

Qy  618 AGCACTACGTGGCGGCGCCGCCGCGCATGTTGGGGGCGCCGCTGCGGCAGCGGCGCGTGC 677
        |||||||||||||||| |||| ||| ||| || ||| ||||| ||||||| ||||||||
Db  478 AGCACTACGTGGCGGCGCCGCCGCGCATGCTGGGGGCGCCGCTGCGGCAGCGGCGCGTGC 537

Qy  678 GGCCGCTGCAGGAGCTGTGCCGCCAGCGCATCGTGGCCACCGTGGGGCGCGAGAACCTGG 737
        ||||||||||||| |||||||||||||||||||||||||||||||||||| ||||||||
Db  538 GGCCGCTGCAGGAGCTGTGCCGCCAGCGCATCGTGGCCACCGTGGGGCGCGAGAACCTGG 597
```

```
Qy  492 ACTGCTTCTTCGGCGTCTTCAGCGTGACGATGGGTCTTCGGCGGCCCACGAGGCATCCGGCGTGCCACT 551
        ||||| || |||||  || |||||| || ||||| | |||| || ||| |||||||||| ||||||||||
Db  483 ACTGCTTTTCGCCCTTAGCGTGAGAGATGGCCTCGGGACCACGACGCATCCGCGTGCCACT 542

Qy  552 TCCAGGCCGGCGCGCTTCCACTTGGACGGCCAGCGCCGCGAGAGACTTCGACTGCCTTTTCGGAGC 611
        | ||| || || |||| || || || |||| ||| || ||| ||||||| |||||| |||| ||
Db  543 TTCAGCGCCGCTTCACCTTGGATGGCAGGCGCGGAGAGCTTCGACTGCCTCTTCGAGC 602

Qy  612 GCGTCGGCGCGGCCCCCGCCCCGCCCCGGGCCCCGCGCTGGGCCGGCGCCGCCTGGGCCCACGCGCC 671
        || | ||  |||| |||| ||||| |||| || |||||||| || || |||||||| |||| ||||
Db  603 GCGGTGGCGCACTACGTGGCGGCCCCGCCGCCCCGCCGCCTGGGGCGGCC 662

Qy  663 GCGGTGGCGGGCCCCTGCAGGAGCTGTGCCGCCAGCATCGTGGCCACCGTGGGGGTCGCCCCAGCGCC 722
        || ||||| |||| || ||||||| |||||||||| |||| |||| ||| || || |||| ||||||
Db  543 ... (continued)

Qy  672 GCGGTCGGCGGGCCCCTGCAGGAGCTGTGCCGCCAGCGCATCGTGGCCACCGTGGGGGTCGCGGG 731
        |||| | ||||| ||| |||||| ||||||||||| ||||||| |||| |||| || || |||||
Db  603 ...

Qy  732 ACCTGGCGCGGGCCATCCCTCTTAACCCGGGTACTCCGGTGACTACCGGAGTTCCTTCCCCTTCC 791
        ||| || || |||||| || |||  ||||||| || |||||| ||||| || |||||| ||||
Db  723 ACCTGGCCATCCCCAAACCCGTCCTCCGGGACTACCTGAGCTCCTTCCCTTTCC 782

Qy  792 AGATCTGACCGGCTGCCCGCTGTGCCCAGCATTAAGTGGGGGGGCGGCCTTATTATTTCTTAT 851
        || ||||| || ||| || ||||||| || |||  || ||||| || ||||| ||||| |||
Db  783 AGATTGACCGCGGCAGCCCCCGTCGCACGCAGCATTAACTGGGATGCCCGGTAGGCAGCATTAACTGGAT 840

Qy  852 TATTAATAATAATTATTTCTCGGAACCACCTGGGATGGGAGGGACCCCTCCC-GCCTCGGGTCGTGTTATT- 910
        | ||||| || ||||| | || || |||||||||| ||||| ||| |||  |||| ||||| | |||
Db  852 TATTAATAATAATTATTTCTCGGAACCACCTGGGATGGGAGGGACCCCTCCC-GCCTCGGGTCGTGTTATT- 910

Qy  911 GAGTGGGTTGTG---GAGGGTGAGAGAATGGCCTCCCCACTTCTGGCCTGGAGA------CCTCATCC 962
        |||||||||||   ||||| |||| ||||| || ||| ||||| |||| ||||      |||| ||
Db  894 GAGCGGATGGGGTGTAGGGACGGAGGCCCTCGGCCTCGGCTGGAGGACGAGCGCCCGCAGAC 953

Qy  963 CACCCTCAGGGGTGGGGTGGTGCT-----CCCTCCTCTGGTGTCCCTCCGGGTCCCCCC--CT 1016
        ||| ||||||| ||||||||| ||     ||||| ||| ||||| |||| ||||| |||  ||
Db  954 CCCTTCTCACCTCTTGAGCGGGGTCTCCCCGTCTCCCTTGTGTGCTTCCCCTCCCCTGGT 1013

Qy 1017 GGTTGGTTAGCACGCT------TGTGTCTCGGGGGCCAGGAACCTGAATT--CCACTCCTTAACCTCTCC 1070
        || |||||||| ||||      |||| || ||||| ||||| ||| |||  |||| | |||| |||||
Db 1014 TGTTGGTTAGCACTTAACTGTAATCTGAGCCGACAGGCCCTCCTACCTCTTC 1073

Qy 1071 ATGTTTACATATNTCCCAGTATCTTTGCACAAACAGG---TCGGGGAGGGGTGTCTCGGCTT 1129
        || ||| |||| |||||||| || || |||  || || |||| ||| || |||| |||| ||
Db 1074 ATGTTTACATATACCCAGTATCTTGCACAAAACCAGGGGTTGGGGAGGGGTCTCGGCTT 1133

Qy 1130 CATTTTTTTCTCTGTGTGCAGAANTATCTTATTTTTAATTTTAATTTTACAGCCAGTTTAGGTAATAAA 1189
        ||| ||||| |||||| |||| || ||| | |||||| ||||| |||||||| || ||||| || |||
Db 1134 TATTTTTCTCGTGTGCGAGAAATCCTATTTTTAITAT-TTTTTAAAGTCAGTTTAGGTAATAAA 1192

Qy 1190 CTTTATTATGAAAAGTTTTTTTTTT 1212
        |||||||||||| ||||||||||||
Db 1193 CTTTATTATGAAAAGTTTTTTTTTT 1215
```

```
Qy     738 CGCGCATCCCTCTTAACCCGGTACTCCGTGACTACCTGAGTTCCTTCCCCTTCCAGATCT 797
           | |||||||| || ||||| || ||||| ||||||||||| ||||||||||||||||| |
Db     598 CTCGCATCCCCCTCAACCCCGTCCTCCGCGACTACCTGAGCTCCTTCCCCTTCCAGATTT 657

Qy     798 GACCGGCTGCCGCTGTGCCGCAGCATTAAGTGGGGGCGCCTTATTATTTCTTATTATTAA 857
           ||||||| ||  | |    |||      | |    | |      | ||  |
Db     658 GACCGGCAGCGCCCGCCGTGCACGCAGCATTAACTGGGATGCCGTGTTATT--------- 709

Qy     858 TTATTATTATTTTTCTGGAACCACGTGGGAGCCCTCCCC-GCCTGGGTCGGAGGGAGTGG 916
           || |||||| ||  ||||||||| |||||  |||||||| |||||||| |||||||| ||
Db     709 TTGTTATTACTTGCCTGGAACCATGTGGGTACCCTCCCCGGCCTGGGTTGGAGGGAGCGG 768

Qy     917 TTGTG---GAGGGTGAGATGCCTCCCACTTCTGGCTGGAGA-----CCTCATCCCACCTC 968
            || |     ||| |||  ||||||| |    |||||||||     || ||  || | ||
Db     769 ATGGGTGTAGGGGCGAGGCGCCTCCCGCCCTCGGCTGGAGACGAGGCCGCAGACCCCTTC 828

Qy     969 TCAGGGGTGGGGGTGCT----CCCCTCCTGGTGCTCCCTCCGGGTCCCCC--CTGGTTGT 1022
           |||    | | || | |    ||||||||||||||||||| |||||||||   | |||||
Db     829 TCACCTCTTGAGGGGGTCCTCCCCCTCCTGGTGCTCCCTCTGGGTCCCCCTGGTTGTTGT 888

Qy    1023 AGCAGCT----TGTGTCTGGGGCCAGGACCTGAATT--CCACTCCTACCTCTCCATGTTT 1076
           |||||||    ||| ||||| ||||||||||||| |  |  ||||||||||| |||||||
Db     889 AGCAGCTTAACTGTATCTGGAGCCAGGACCTGAACTCGCACCTCCTACCTCTTCATGTTT 948

Qy    1077 ACATATTCCCAGTATCTTTGCACAAACCAGGGG-TCGGGGAGGGTCTCTGGCTTCATTTT 1135
           |||||| |||||||||||||||||||||||||| | |||||||||||||||||| |||||
Db     949 ACATATACCCAGTATCTTTGCACAAACCAGGGGTTGGGGGAGGGTCTCTGGCTTTATTTT 1008

Qy    1136 TCTGCTGTGCAGAATATCCTATTTTATATTTTTACAGCCAGTTTAGGTAATAAACTTTAT 1195
           |||||||||||||||    | || ||| |||||| || ||||||||||||||||||||||
Db    1009 TCTGCTGTGCAGAATCCTATTTTATAT-TTTTTAAAGTCAGTTTAGGTAATAAACTTTAT 1067

Qy    1196 TATGAAAGTTTTTTTTTAAAAGAAAAA 1222
           |||||||||||||||||  || |||||
Db    1068 TATGAAAGTTTTTTTTTTTAAAAAAAA 1094

RESULT  13
AB005043
LOCUS       AB005043     1082 bp    mRNA            PRI       27-SEP-1997
DEFINITION  Homo sapiens mRNA for STAT induced STAT inhibitor-1, complete cds.
ACCESSION   AB005043
VERSION     AB005043.1  GI:2443364
KEYWORDS    STAT induced STAT inhibitor-1.
SOURCE      Homo sapiens cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1082)
  AUTHORS   Minamoto,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-1997) to the DDBJ/EMBL/GenBank databases. Seijiro
            Minamoto, Osaka University, Internal Medicine III; Yamadaoka 2-2,
            Suita, Osaka 565, Japan (E-mail:minamoto@imed3.med.osaka-u.ac.jp,
            Tel:06-879-3833, Fax:06-879-3839)
REFERENCE   2  (sites)
  AUTHORS   Minamoto,S., Ikegame,K., Ueno,K., Narazaki,M., Naka,T.,
            Yamamoto,H., Matsumoto,T., Saito,H., Hosoe,S. and Kishimoto,T.
  TITLE     Cloning and functional analysis of new members of STAT induced STAT
            inhibitor (SSI) family : SSI-2 and SSI-3
  JOURNAL   Biochem. Biophys. Res. Comm. 237, 79-83 (1997)
FEATURES             Location/Qualifiers
     source          1. .1082
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             54. .689
                     /note="SSI-1"
                     /codon_start=1
                     /evidence=experimental
                     /product="STAT induced STAT inhibitor-1"
                     /protein_id="BAA22431.1"
                     /db_xref="GI:2443365"
                     /translation="MVAHNQVAADNAVSTAAEPRRRPEPSSSSSSPAAPARPRPCPA
```

```
                     VPAPAPGDTHFRTFRSHADYRRITRASALLDACGFYWGPLSVHGAHERLRAEPVGTFL
                     VRDSRQRNCFFALSVKMASGPTSIRVHFQAGRFHLDGSRESFDCLFELLEHYVAAPRR
                     MLGAPLRQRRVRPLQELCRQRIVATVGRENLARIPLNPVLRDYLSSFPFQI"
BASE COUNT      154 a    396 c    316 g    216 t
ORIGIN

  Query Match             55.3%;  Score 682.4;  DB 9;  Length 1082;
  Best Local Similarity   83.4%;  Pred. No. 8.2e-114;
  Matches  904;  Conservative    0;  Mismatches  146;  Indels   34;  Gaps   10;

Qy     108 CAGCGCAGCCCCGGACGCTATGGCCCACCCCTCCAGCTGGCCCCTCGAGTAGGATGGTAG 167
           ||||||| |||||||||||||||||||||||||| ||||||||||   ||||||||||||
Db       1 CAGCGCACCCCCGGACGCTATGGCCCACCCCTCCGGCTGGCCCCTTCTGTAGGATGGTAG 60

Qy     168 CACGCAACCAGGTGGCAGCCGACAATGCGATCTCCCCGGCAGCAGAGCCCCGACGGCGGT 227
           ||| ||||||||||||||||||||||||  ||||| | |||||||||||||||||||||
Db      61 CACACAACCAGGTGGCAGCCGACAATGCAGTCTCCACAGCAGCAGAGCCCCGACGGCGGC 120

Qy     228 CAGAGCCCTCCTCGTCCTCGTCTTCGTCCTCGCCAGCGGCCCCCGTGCGTCCCCGGCCCT 287
           |||| || |||   ||||| || || |||||||| |||||||||| ||| || ||||| |
Db     121 CAGAACCTTCC---TCCTCTTCCTCCTCCTCGCCCGCGGCCCCCGCGCGCCCGCGGCCGT 177

Qy     288 GCCCGGCGGTCCCAGCCCCAGCCCCTGGCGACACTCACTTCCGCACCTTCCGCTCCCACT 347
           |||| |||||||| ||||| ||||| |||||||| ||||||||||| ||||| || |||
Db     178 GCCCCGCGGTCCCGGCCCCGGCCCCCGGCGACACGCACTTCCGCACATTCCGTTCGCACG 237

Qy     348 CCGATTACCGGCGCATCACGCGGACCAGCGCGCTCCTGGACGCCTGCGGCTTCTATTGGG 407
           ||||||||||||||||||||||  ||||||||||||||||||||||||| ||||| ||||
Db     238 CCGATTACCGGCGCATCACGCGCGCCAGCGCGCTCCTGGACGCCTGCGGATTCTACTGGG 297

Qy     408 GACCCCTGAGCGTGCACGGGGCGCACGAGCGGCTGCGTGCCGAGCCCGTGGGCACCTTCT 467
           | ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db     298 GGCCCCTGAGCGTGCACGGGGCGCACGAGCGGCTGCGCGCCGAGCCCGTGGGCACCTTCC 357

Qy     468 TGGTGCGCGACAGTCGTCAACGGAACTGCTTCTTCGCGCTCAGCGTGAAGATGGCTTCGG 527
           ||||||||||||| || || ||||||||||| ||||| || |||||||||||||| ||||
Db     358 TGGTGCGCGACAGCCGCCAGCGGAACTGCTTTTTCGCCCTTAGCGTGAAGATGGCCTCGG 417

Qy     528 GCCCCACGAGCATCCGCGTGCACTTCCAGGCCGGCCGCTTCCACTTGGACGGCAGCCGCG 587
           | ||||||||||||||||||||||| |||||||||||||| ||| |||| ||||||||||
Db     418 GACCCACGAGCATCCGCGTGCACTTTCAGGCCGGCCGCTTTCACCTGGATGGCAGCCGCG 477

Qy     588 AGACCTTCGACTGCCTTTTCGAGCTGCTGGAGCACTACGTGGCGGCGCCGCGCCGCATGT 647
           ||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     478 AGAGCTTCGACTGCCTCTTCGAGCTGCTGGAGCACTACGTGGCGGCGCCGCGCCGCATGC 537

Qy     648 TGGGGGCCCCGCTGCGCCAGCGCCGCGTGCGGCCGCTGCAGGAGCTGTGTCGCCAGCGCA 707
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     538 TGGGGGCCCCGCTGCGCCAGCGCCGCGTGCGGCCGCTGCAGGAGCTGTGCCGCCAGCGCA 597

Qy     708 TCGTGGCCGCCGTGGGTCGCGAGAACCTGGCGCGCATCCCTCTTAACCCGGTACTCCGTG 767
           |||||||| ||||||| |||||||||||||| |||||||| || ||||| || ||||| |
Db     598 TCGTGGCCACCGTGGGCCGCGAGAACCTGGCTCGCATCCCCCTCAACCCCGTCCTCCGCG 657

Qy     768 ACTACCTGAGTTCCTTCCCCTTCCAGATCTGACCGGCTGCCGCTGTGCCGCAGCATTAAG 827
           |||||||||| ||||||||||||||||| |||||||| ||  | |    |||      |
Db     658 ACTACCTGAGCTCCTTCCCCTTCCAGATTTGACCGGCAGCGCCCGCCGTGCACGCAGCAT 717

Qy     828 TGGGGGCGCCTTATTATTTCTTATTATTAATTATTATTATTTTTCTGGAACCACGTGGGA 887
           |    | |      | ||  |         || |||||| ||  ||||||||| |||||
Db     718 TAACTGGGATGCCGTGTTATT---------TTGTTATTACTTGCCTGGAACCATGTGGGT 768

Qy     888 GCCCTCCCC-GCCTGGGTCGGAGGGAGTGGTTGTG---GAGGGTGAGATGCCTCCCACTT 943
            |||||||| |||||||| |||||||| || || |     ||| |||  ||||||| |
Db     769 ACCCTCCCCGGCCTGGGTTGGAGGGAGCGGATGGGTGTAGGGGCGAGGCGCCTCCCGCCC 828

Qy     944 CTGGCTGGAGA-----CCTCATCCCACCTCTCAGGGGTGGGGGTGCT----CCCCTCCTG 994
             |||||||||     || ||  || | |||||    | | || | |    |||||||||
Db     829 TCGGCTGGAGACGAGGCCGCAGACCCCTTCTCACCTCTTGAGGGGGTCCTCCCCCTCCTG 888

Qy     995 GTGCTCCCTCCGGGTCCCCC--CTGGTTGTAGCAGCT----TGTGTCTGGGGCCAGGACC 1048
           |||||||||| |||||||||   | ||||||||||||    ||| ||||| |||||||||
```

```
Db   889  GTGCTCCCTCTGGGTCCCCCTGGTTGTTGTTGTAGCAGCTTAACTGTATCTGGAGCAGGGACC  948

Qy  1049  TGAATT--CCACTCCTACTCTCCATGTTTACATATTCCCAGTAGTACTTTGCACAAACCAG  1106
          |||||| |||||||||||||  ||||||||| |||||||||||||||||||||||||||||||
Db   949  TGAACTCGCACCTCCTACCCTCTTCCATGTTTACATATACCCAGTATCTTTGCACAAACCAG  1008

Qy  1107  GGG-TCGGGGAGGGGTTCTCGGCTTCATTTTTTCTGCGTGTGCAGAAATATCCTATTTTATATT  1165
          ||| ||||||||||||||||||||||||||||||||||||||||||||| ||| |||||||||
Db  1009  GGGGTTGGGGAGGGGTTCTCGGCTTTATTTTTTCTGCGTGTGCAGAAACCTATTTTATATTT  1068

Qy  1166  TTTA  1169
          |||
Db  1069  TTAA  1072
```

```
Qy   207  CAGCAGAGAGCCCCGACGGACGGGGGTCAGAGCCCTCCTCTCCTTCGTCTTCGCTCGCCAGCGG  266
          |||||||||||| ||||||||||||||| ||||  |   | ||||||||||||||| |||||||
Db     5  CAGCAGAGCCCCGACGGGCGGCCAGAAGCCTTCC---TCCTCTTCCTCCTCCTCGCCCGCGG  61

Qy   267  CCCCCGCGTGCGGTCTCCCGGCCCTGCCCCGCGGCCCTGCCCCCCGGGCGACACTCACT  326
          ||||| ||  || ||||||||||||||||||||| |||||||| ||||||||||||||||
Db    62  CCCCCGCGGCCCGCCGGGCGGGCCCGCCCCGCGGGCCTGCCCCCCGGGCGGACACGCACT  121
```

```
Qy   327  TCCGCACTTCCGCTCCCCACTCCGATTACCGGGCGCATCACGGGCGCCAGCGGACCAGCGGG  386
          ||||| |||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Db   122  TCCGCACATTCCGGTTCGCGACCCCGATTACCGGCGCATCACGGGCGCCAGCGCCGCCTCGGG  181

Qy   387  ACGCCTGCGGCGTTCTTCATTGGGGAGCCCCTGACCGTGACCGTGGGCGCCACGGACGGGGG  446
          ||||||||| |||||||||||||||||||||| ||||||||||||||||||| |||||||||
Db   182  ACGCCTGCGGGATTCTCACTGGGGGCCCCTGAGGCGTGACCGTGCGCACGGAGCGGCTGCCG  241

Qy   447  CCGAGCCCGTGGGCCACCTTCTTGGGTCGGGCGGACAGTCGTCAACGGGGAACTGCTCTCGGC  506
          ||||||||| ||||||||||||||||||||||||||||||| |||||||||| ||| ||||||
Db   242  CCGAGCCCGTGGGCACCTTCCTGGGTCGGGCGGACAGTCGCCAGCGGAACTGCTTTCCGC  301

Qy   507  TCAGCGTGAAGAGATGGCTTCGGCGCCCCACGAGCGCATCCGCGTGCGACTTCCAGGGCCGG  566
          |||||||| |||||||||||||||||| ||||||||||||||| |||||||||||||||||
Db   302  TTAGCGTGAAGAGATGGCTTCGGCGCCCCACGAGCGCATCCGCGTGCGACTTCCAGGGCCGCT  361

Qy   567  TCCCACTTGGAGGGGCCACGGCGCGGCAACCTTCGGCTGCTTTTGCGGGTGCTGCGAGCACTACG  626
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   362  TTCACCTTGGATGGGCAGGCCGGGGAGGCTTCGGCTGCCTTCGGGTGCTGGGAGCACTACG  421

Qy   687  AGGAGCTGTGGCCCAGGCATCGTGGGGGGGCCGCGGGTCGGCGAGAACCTGGGGCCGGGCCATCC  746
          ||| |||||||| |||||||  ||||||||||||||||||||||||| |||||||||||| ||||
Db   422  TGGGCGCGCCCGGCCCCCATGCTGCGGGGGCCCCGCGGGTCGGGCGCGGGACAGTCGTCAACGGG  481

Qy   482  AAGAGCTGTGGCCCAGGCATCGTGGCCACGTGGGGTCGGCGAGAACCTGGCTCGCATCC  541

Qy   747  CTCTTAAACCCGGTACTCCCGTGACTACCTGAGTTCCTTCCCCCTTCCGAGATCCGACCCGGCTG  806
          ||||| |||||||| |||| |||||||||||||||||||||||||||| |||||||||||||||
Db   542  CCCTCAACCCCGTCCGCGACTACCTGAGCGTCCTCCCCCTTCCCGAGATTTGGACCGGCAG  601

Qy   807  CCGGCTGTGCCCAGCAGCATTAAGTGGGGGCGGCGGCCTTATATATTTCCTTATTATTAATTATTA  866
          ||||||||||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||
Db   602  CGCCCCCGTGCCCAGCGCAATTAAGTGGGGGATGCGGGTTGTTATT------TTGTTATTA  652

Qy   867  TTTTTCTGGAACCACGTGGGAGCCCCTCCCC--GCCTGGGGTCGGGGAGGGGAGTGGTTGTG---G  922
          ||||| | ||||||||||||||||||||||   ||||||||||||||||||||||||||||   G
Db   653  CTTGCCTGGAACCCATGGGTACCCTCCCGCCCTGGGTTGGGGAGGGGAGGGGAGTGGGGGTA  712

Qy   923  AGGGTGAGATGCCTCCCACTTCTGGGTGGGAGA------CCTCATCCACCTCTCAGGGGTG  977
          ||||||||| | ||||| ||| |||| ||||||      ||| ||| ||||| |||||||
Db   713  GGGGGCGGAGGCGGCCTCCCGGCCTGGAGACGAAGCGCAGAGACCCCTTCTCACCCTCTT  772

Qy   978  GGGGTGGCT-----CCCTCCGGTGCCTCCCTCCGGGGTCCCCC--CTGGTTGTAGGCAGCTT-  1031
          |||||||||     ||||| ||||||||||||||||||||||   |  || | ||| ||||
Db   773  GAGGGGGGTCCTCCCCCTCCTGGTGCCCCCTCTGGGGTCCCCCTGGGGTTGTTGTTGAGCAGCTTA  832

Qy  1031  ---GTGCTGGGGGCCAGGAACCTGAATT--CCACTCCTACCTCCCATGTTGTTACATATTCC  1085
             |||||||||||||||||||||||| |  ||||||||||| |||||||||||||||| |||
Db   833  AGTGTATCTGGAGCGCAGGACCTGAACTCGCACTCCTACCTCTTCCATGTTTACATATACC  892

Qy  1086  CAGTAGTTTGCACAAACCAGGGGTGCGGGAGGGGTCTCGTGGC-TTCATTTTTTCTGCGTGTG  1144
          ||||| |||||||||||||| |||| || ||||||| ||| ||| ||| |||||||||||||
Db   893  CAGTATCTTTGCACAAACCATGGTTTGGGGAGGTCTCCGGGGTTTTATTTTTTCTGCGTGTG  952

Qy  1145  CAGAAATATCCTATTTTATATTTTTATATTTTTACAGCCAGTTTAGGTAATAAACTTTATATGAAAGT  1204
          ||||||  | |  ||||||||||||||| ||||||||||||||||||| |||||||||||| |||
Db   953  CAAAAACCCTAATTTTTATATTTTTTTAAATNCANTTTAGGTAATAAACTTTATTAATAAAAT  1012

Qy  1205  TTTTTTTTTAAAGAAAAA  1222
          |||||||| || ||||
Db  1013  TTTTTTTTTAAAAAAA  1030
```

```
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 908)
AUTHORS    Yandava,C.N., Pillari,A. and Drazen,J.M.
TITLE      Radiation hybrid and cytogenetic mapping of SOCS1 and SOCS2 to
           chromosomes 16p13 and 12q, respectively
JOURNAL    Genomics 61 (1), 108-111 (1999)
MEDLINE    99443878
REFERENCE  2  (bases 1 to 908)
AUTHORS    Yandava,C.N., Pillari,A. and Drazen,J.M.
TITLE      Direct Submission
JOURNAL    Submitted (01-MAR-1999) Pulmonary and Critical Care Medicine,
           Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115,
           USA
FEATURES             Location/Qualifiers
     source          1..908
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     mRNA            <14..>649
                     /product="suppressor of cytokine signalling-1 SOCS-1"
     CDS             14..649
                     /codon_start=1
                     /product="suppressor of cytokine signalling-1 SOCS-1"
                     /evidence=not_experimental
                     /product="suppressor of cytokine signalling-1 SOCS-1"
                     /protein_id="AAD27709.1"
                     /db_xref="GI:4680639"
                     /translation="MVAINNQVAADNAVSTAAEPRRPEPSSSSSSSPAAPARPRPCPA
                     VPAPAPGDIHFRTFRSHADYRRITRASALLDACGFYWGPLSVHGAHERLRAEPVGTFL
                     VRDSRQRNCFFALSVKMASGPTSIRVHFQAGRFHLDGSRESFDCLFELLEHYVAAPRR
                     MLGAPLRQRRVRPLQELCRQRIVATVGRENLARIPLNPVLRDYISSFPFQI"
BASE COUNT       120 a    347 c    280 g    161 t
ORIGIN
```

Query Match
Best Local Similarity   46.3%;
                        82.6%;   Pred. No. 7.5e-94;
Matches 760; Conservative 0; Mismatches 129; Indels 31; Gaps 8;
Score 571.6; DB 40; Length 908;

```
Qy   148 CCCCTCGAGTAGGATAGGATGGTAGCACCGCAACCACCAGGTGGCAGCGGGATCTCCCCGGC 207
         |||||      |||||||| ||||| ||||| ||||||||||||||||| |||||| ||| ||
Db     1 CCCCTTCTGTAGGATGGGTAGCACACAACCAGGTGGCAGCGCGACAATGCCGATCTCCACAGC 60

Qy   208 AGCAGAGCCCGACGGGCGGTCAGAGCCCTCCTCGTCTCGTCTTCGTCCTTCGCCAGCGGC 267
         ||||||||| ||| | ||||| | |||  || || || | |||||| |||| ||||  |||
Db    61 AGCAGAGCCCCGAGCGGCGGCCAGAACCTTCC---TCCTCTTCCTTCCTTCTCCGCCGGC 117

Qy   268 CCCCGTGGGTCCCCGGCCCTGCCCGGCGGTCCCAGCCCCTTGCCGCGACTCACTT 327
         ||||| ||||||||| ||||  |||  |||  ||||| | |  ||||||||||||
Db   118 CCCGCGCCCCGGGCGGTCCCCGGCGTCCCCGGGGACACGACACTT 177

Qy   328 CCCGCACCTTCCGCTCCCACTCCGATTACGGGCCGATCACGGCAGGGGGTCCTCGGA 387
         ||||||||| |||||||||| ||| ||||| |||  |||||||||||| |||||||
Db   178 CCCGCAATTCCGTTCGCACGCCGATTACCGCGCGCCCGCCCGGGGGACACGCACTT 237

Qy   388 CGCCTGCGCGCCTTCTATTGGGGACCCCTGAGCGTCGCACGGACGGGCTGCGTGC 447
         ||||||  |||| | || |||||| ||||| |||| | ||| |||||| | ||
Db   238 CGCCTGCCGGAATTCTACTGGGGGCCCCTGAGCGTTGCACGGGGGCTGCGC 297

Qy   448 CGAGCCCGTGGGGCACCTTCTTGTGGTGGCGGGGACCGGGATCTGAAACGAGAACTGCGC 507
         ||||||||  ||| |||| |   || || | |||  | | |||| | ||||||||| |
Db   298 CGACCCCGTGGGCCACCTTCCTGGTTGGCCGGCCAGCTGTTTTTCGCCCT 357

Qy   508 CAGGGTGAAGAGATGGCTTCGGGCGTGGTTGGCGCCACTTCCAGGCCGCGGCTT 567
         || | |||||||||||| |||||||||| |||||||||||| |||||||  |
Db   358 TACGGTGAAGATGGCCTCGGGGACCACCATCCGCTGCACTTTCAGGCCGCGCTT 417

Qy   568 CCACTTGGACGGCAGCCGCGGAGACCTTCGACTGCCTGCCTTGCTGGAGCACTACGT 627
         || ||||| |||||| ||| |||||||| |||||||||| | ||||||| |||  ||
Db   418 TCACCTGGATGGCAGCGGCGGAGACTTCGACTGCCTCTTCGAGCTGCTGGAGCACTACGT 477

Qy   628 GGCGGCGCCGCGCGGCCCGCCGCCATGTTGGGGCGGGCCCCGGCCCTGCGGCAGCGCCGTGCA 687
         |||| || || ||| ||| | |||||||||||||||||||| |||||||||||| ||||||
Db   478 GGCGGCGCCGCGCGGCCCGCCGCCCATGCTGGGGCGGGCCCCGCTCCAGGAGCTGTGCCGTCAGCGCA 537

Qy   688 GGAGCGTGTGCGCCAGCGCGATCGTGGGGGCGGGCCCCGCGTCCGGAGAACCTGGCCGCCATCCC 747
         |||| | ||| | |||| | || ||||||||| | |||||  |||||||||||| |||||||
Db   538 GGAGCTGTGCCGCCAGCGCATCGTGGCCACCGTGGGCCGGGAGAACCTGGCTCGCATCCC 597

Qy   748 TCTTAACCCGGTGACTTCCGGTACTACCTGAGTTCCTTCCCCTTCCAGAGATCTGACCGGGCTGC 807
         |  |||  | || || |  ||||| |||   || || |||||  ||||| | ||||| | |||
Db   598 CCTCAACCCCGTCCTCCGCGACTACCTGAGCTCCTTCCCCTTCCAGATCTGACCGGGCAGC 657

Qy   808 CGCTGTGCCCGCAGCAATTAAGTGGGGCGGGGCGCCTTATTATTTCTTATTATTAATTATTATTAT 867
         ||| |     |||||  | ||| |  |||  ||||||| |  || | |  ||  || | |
Db   658 GCCCCCGCGCTGCACGCAGCAATTAACTGGGATGCCGGTGTTATT---------TTGTTATTAC 708

Qy   868 TTTTCTGGAACCACGTGGGAGAGCCCTTCCCC--GCCTGGGGTCGGGAGGGAGGGAGGTGGTTGTTG---GA 923
         |  |||| |||| || ||| || |||||||||  |||  |||| |||||| ||| |  |     | | ||
Db   709 TTGCCTCGGACCATGTGGGTACCCCTCCCCCGGGTTGGGTTGGAGGGGCGGGGATTGGGGTGTAG 768

Qy   924 GGGTGAGATGCCTCCCACTTCTGGGCTGGAGA----CCTCATCCCACCTCTCAGGGGTGG 978
         ||| |||||| ||||||| | ||||||||||     || |  |  || | ||||| |
Db   769 GGGCGAGGGGCGCTCCCGGCTGGAGACGAGGACCCTTCCGGGAGGGGTTGTAGCAGCT 828

Qy   979 GGGTGCT----CCCTTCCTGGTGCTCCCCTCCCGGGGTCCCCCC--CTGGTTGTTGAGCAGCT--- 1030
         ||| |||    |||||||||||||||||| | ||  |||||||   | ||||||| ||||||
Db   829 AGGGGGGTCCTCCCCCTCTCGGGGTCCCCCCTCTGGGTGTTGTGTTGTGAGCACTTAA 888

Qy  1030 -TGTGTCTGGGGCGCAGGACC 1048
          |  | |||||||| |||||
Db   889 CTGTATCTGAGGCCAGGAGGACC 908
```